Behavioral/Systems/Cognitive

Gamma-Band Activity as a Signature for Cross-Modal Priming of Auditory Object Recognition by Active Haptic Exploration

Till R. Schneider, Simone Lorenz, Daniel Senkowski, and Andreas K. Engel

Department of Neurophysiology and Pathophysiology, Center of Experimental Medicine, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany

When visual sensory information is restricted, we often rely on haptic and auditory information to recognize objects. Here we examined how haptic exploration of familiar objects affects neural processing of subsequently presented sounds of objects. Recent studies indicated that oscillatory responses, in particular in the gamma band (30–100 Hz), reflect cross-modal processing, but it is not clear which cortical networks are involved. In this high-density EEG study, we measured gamma-band activity (GBA) in humans performing a haptic-to-auditory priming paradigm. Haptic stimuli served as primes, and sounds of objects as targets. Haptic and auditory stimuli were either semantically congruent or incongruent, and participants were asked to categorize the objects represented by the sounds. Response times were shorter for semantically congruent compared with semantically incongruent inputs. This haptic-to-auditory priming effect was associated with enhanced total power GBA (250–350 ms) for semantically congruent inputs and additional effects of semantic congruency on evoked GBA (50–100 ms). Source reconstruction of total GBA using linear beamforming revealed effects of semantic congruency in the left lateral temporal lobe, possibly reflecting matching of information across modalities. For semantically incongruent inputs, total GBA was enhanced in middle frontal cortices, possibly indicating the processing or detection of conflicting information. Our findings demonstrate that semantic priming by haptic object exploration affects processing of auditory inputs in the lateral temporal lobe and suggest an important role of oscillatory activity for multisensory processing.

Introduction

When grasping an object in the dark, for instance, when searching for the alarm clock at the bedside in the morning, we strongly rely on our tactile and auditory senses. Haptic information is provided via the tactile and proprioceptive sensory systems stimulated by touching objects. The physical features of haptically perceived objects like roughness, hardness, or weight differ substantially from the features of the sounds of objects such as tone, pitch, loudness, and spectral composition (Lederman and Klatzky, 1987, 1993). Haptic information may thus facilitate auditory object recognition by providing complementary information about the object identity. Several studies reported evidence for cross-modal priming between the visual and the haptic systems (Easton et al., 1997; Reales and Ballesteros, 1999; James et al., 2002). On the cortical-level, lateral occipitotemporal areas have been impli-

cated in the conjoint processing of visual and haptic object-related information (Amedi et al., 2001; Peltier et al., 2007). Moreover, inferior frontal gyrus and insula have been shown to respond to both auditory and haptic object identity (Renier et al., 2009). Yet, to our knowledge, it has not been investigated whether haptic information can directly facilitate auditory object recognition.

In animal studies, multisensory interactions between the somatosensory and the auditory system have been observed in the caudal-medial belt area of the auditory cortex in macaque monkeys (Schroeder et al., 2001; Fu et al., 2003). Additionally, multisensory integration of touch and sound occurs in the caudal-medial belt area, which is in close proximity to primary sensory auditory cortex, as has been found using high-resolution functional magnetic resonance imaging of the macaque monkey (Kayser et al., 2005, 2009).

Although the presence of multisensory interactions at the single-neuron level in various cortical areas is well established, the underlying neural mechanisms are much less well understood. Within sensory modalities, highly specific patterns of synchronous neural firing may serve as a mechanism for flexible feature integration (Engel et al., 1992, 2001; Singer and Gray, 1995). Moreover, we have recently provided evidence that cross-modal priming due to semantically congruent visual-auditory inputs is reflected by an enhancement of gamma-band activity (GBA; > 30 Hz) in the middle temporal gyrus (MTG) (Schneider et al., 2008b). This suggests that synchronized oscillatory activity may constitute a crucial mechanism for cross-modal processing
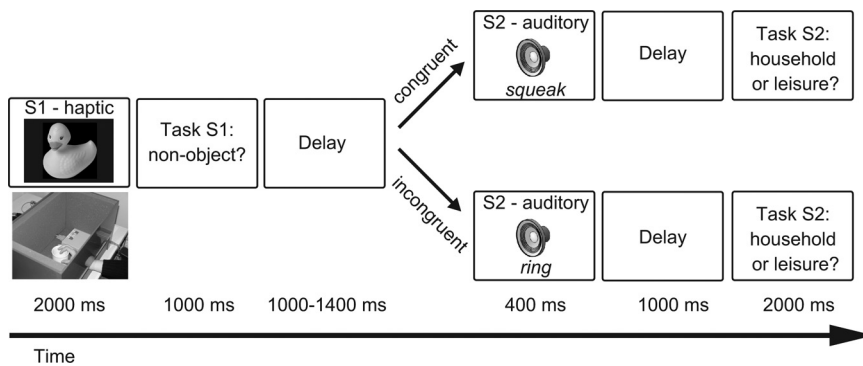
Figure 1. Schematic representation of the cross-modal priming paradigm. The active exploration of S1 was limited to 2000 ms, as marked by acoustic signals. S2 values were either semantically congruent or incongruent to those of S1. To segregate activity related to the motor response from the activity of interest related to the auditory stimulus, the response of the participants was prompted following a delay of 1000 ms after the presentation of S2.

(Kaiser et al., 2005; Lakatos et al., 2007; for review, see Senkowski et al., 2008; Kayser et al., 2008; Maier et al., 2008).

The central goals of the present study were (1) to investigate cross-modal semantic priming between the haptic and the auditory modalities and (2) to examine the modulation of GBA due to semantically congruent and incongruent inputs in these modalities. High-density EEG recordings were obtained from the human scalp, and linear beamforming was applied to estimate the cortical sources reflecting the influence of active haptic object exploration on the recognition of subsequently presented sounds. We predicted that (1) haptic object exploration facilitates auditory object recognition of semantically congruent stimuli, and (2) that this cross-modal priming is reflected in modulations of GBA in areas of multisensory convergence.

Materials and Methods

Participants
Sixteen right-handed volunteers (6 women; age range, 20–29 years) participated in the study. All participants had normal hearing (hearing loss: < 30 dB), reported no history of neurological or psychic illness, and were paid for participation. Due to extensive eye movement and muscle artifacts, the data of two participants had to be excluded from the analysis.

Stimuli
Haptic stimuli. Twenty-eight objects of the categories "household" and "leisure" (14 from each category) were selected as haptic stimuli for the present experiment. In addition, 10 haptic stimuli were developed that had not been identified as common objects in a previous pilot study. These meaningless stimuli, termed "non-objects," served as target stimuli in the present study. During the experiment, the objects were placed in a custom-made plastic box (length: 31 cm; width: 51 cm; height: 30 cm). Through an opening at the top of the box (length: 10 cm; width: 35 cm), the experimenter first put the objects—one at a time—into the box. Through an opening in the front (length: 10 cm; width: 35 cm), participants then haptically explored the objects with their right hand (Fig. 1).

Auditory stimuli. The majority (22 of 28) of auditory stimuli of the two categories household and leisure were taken from a recently published multimodal stimulus set (Schneider et al., 2008a). Six additional stimuli were taken from an on-line sound database (http://www.freesound.org). All sound files were matched in length (400 ms). The intensity was adjusted across stimuli by equalizing the root mean square power of all sound files. To avoid onset and offset clicking noises, the sound files were filtered, resulting in a 10 ms rise and fall time. Auditory stimuli (22 kHz, 16 bit, mono) were presented to the subjects at 70 dB sound pressure level using Eartone foam-protected air tube earphones (AeroCompany).

Procedure
Learning phase. To avoid learning effects in the experiment, participants had to learn the identity of the haptic and the auditory stimuli in a separate session before the beginning of the experiment. Haptic and auditory object stimuli were presented in two separate blocks, and participants were instructed to name and categorize them. Meaningless haptic stimuli had to be categorized as non-objects. To avoid having participants recognize the haptic objects visually, they were blindfolded with an eye mask. Before and during the haptic exploration phase, auditory white noise was presented via headphones to mask potential acoustic signals of the haptic stimuli, which would provide complementary information about the identity of the object. The presentation of stimuli was repeated until participants could identify and categorize all objects correctly without much effort, which lasted ~30 min.

Experimental phase. Each trial comprised a haptic prime (S1) and a subsequent auditory target stimulus (S2), which were either semantically congruent (45%) or semantically incongruent (45%) (Fig. 1). In 10% of all trials, nonobjects were presented in the haptic modality. The experimenter started the trials after placing the haptic object into the exploration box. Participants were allowed to explore the haptic object (S1) for ~2 s after the presentation of an acoustic signal (sinus tone, 650 Hz, 50 ms). During the exploration phase, auditory white noise was presented via the headphones. The end of the exploration phase was marked by a second acoustic signal (50 ms sinus tone). Subsequently, participants were given 1 s to indicate by button press the exploration of a nonobject; no response was required after the presentation of a meaningful object. The purpose of this task was to ensure that participants actively attended and explored the haptic stimuli. Following a delay interval ranging between 1000 and 1400 ms, the auditory target (S2) was presented for 400 ms. The participants' task on the S2 was to decide as quickly and accurately as possible whether the object belonged to the category household or leisure. To indicate their decision, participants had to press a left or right button with their middle and index finger (mapping counterbalanced over participants). To avoid response-related activity in the stimulus period, participants were instructed to withhold their response until an acoustic signal (50 ms sinus tone) was presented 1000 ms after S2 offset. Importantly, the two tasks on S1 and S2 were independent. Thus, possible priming effects from the haptic object on the processing of the auditory inputs would be implicit. The trial order was individually randomized, presenting each auditory stimulus equally often in combination with a congruent and an incongruent haptic stimulus. Each experimental session consisted of six blocks, each comprising 66 trials. The delivery of the stimuli and the recording of responses were performed using Presentation 9.90 (Neurobehavioral Systems).

To examine whether effects of congruency are specific for multisensory processing or merely reflect amodal (i.e., modality independent) semantic processing, an additional unisensory auditory S1-S2 control experiment was conducted. Similar to the cross-modal main experiment, semantically related and unrelated stimuli were presented as S1 and S2 (for details, see supplemental material, available at www.jneurosci.org as supplemental material).

EEG recordings
High-density EEG recordings were acquired using an active electrode system (EASYCAP) with 124 scalp electrodes mounted into an elastic cap and two additional electro-oculogram channels below the eyes. The data were recorded with a passband of 0.016–250 Hz and digitized with a sampling rate of 1000 Hz using a BrainAmp amplifier system (Brain-Products). EEG data were recorded against a nose tip reference but were re-referenced to common average before subsequent analysis steps. Analysis of the EEG data was performed using Matlab 7.3.0 (Math-Works), EEGLAB 5.03 (http://www.sccn.ucsd.edu/eeglab) (Delorme and Makeig, 2004), and FieldTrip (http://www.ru.nl/fcdonders/fieldtrip). For the off-line analysis, the data were bandpass filtered [(0.3–110 Hz for

time-frequency analysis; 0.3–30 Hz for event-related potentials (ERPs)] and downsampled to 250 Hz. Epochs containing nonstereotyped artifacts (e.g., cable movement, swallowing) were removed. An independent component analysis approach was applied to further reduce artifacts such as eyeblinks, horizontal eye movements, or electrocardiographic activity. Independent components representing artifacts were removed from the EEG data by back-projecting all but these components (for details, see Schneider et al., 2008b). Finally, all trials that still exceeded a threshold of 100 μV were rejected automatically. On average, 4.7% (range: 0.3–24.2%) of all trials in each subject were removed by this thresholding procedure.

Data analysis

Analysis of event-related potentials. For the analysis of ERPs, data epochs were extracted (−300 to 1000 ms, around auditory stimulus onset) and baseline corrected (−300 to 0 ms). Electrophysiological responses during the haptic exploration interval were not analyzed as the exploration onset varied over trials in the range of several tens of milliseconds. Trials comprising nonobject stimuli were excluded from the data analysis. In a first analysis step, mean global field power (GFP) was computed (Lehmann and Skrandies, 1980). GFP in response to the auditory object stimulus was compared between semantically congruent (i.e., the same object was presented haptically in S1) and semantically incongruent trials (i.e., a different object was presented haptically in S1). In a subsequent step, ERPs were compared between congruent and incongruent stimuli using point-wise running *t* tests. A significant difference in conditions was defined, if at least 20 ms of contiguous data (i.e., 5 consecutive sample points at a sample rate of 250 Hz) met an α criterion of 0.01 (for details of this method, see Guthrie and Buchwald, 1991). Finally, repeated-measures ANOVA was performed with the within-design factors congruency (congruent, incongruent), and region (7 regions). Seven regions of interest (ROIs) including 10 channels each were a priori defined as left frontal, middle frontal, right frontal, left temporal, middle central, right temporal, and middle occipital regions. Time windows for this ANOVA were selected according to the peaks in the GFP averaged across both conditions.

Spectral analyses: evoked power. For the calculation of evoked power, a wavelet approach was applied by convolving the data with a complex Morlet wavelet for frequencies from 20 to 100 Hz (step size, 1 Hz). The transformation to the frequency domain was performed on the average of all trials (i.e., the ERP) separately for each experimental condition. Evoked power reflects the phase-locked part of the signal comparable to the ERP, whereby the latter, due to more restricted low-pass filtering, only reflects phase-locked activity <30 Hz. Additionally, the intertrial coherence (ITC) was calculated by means of averaging the phase of the complex wavelet transform of each single trial. The resulting values range between 0 and 1, indicating randomly distributed phases or perfect phase locking to the stimulus, respectively. To compute the relative signal change of evoked responses, data were normalized with respect to the total power baseline as follows: $P(t,f)_{evoked} = 100 \times (P(t,f)_{evoked} - P(f)_{evoked-baseline})/P(f)_{total-baseline}$. For evoked activity, grand mean time-frequency representations (TFRs) were computed over all participants to illustrate the relative change of activity to baseline. In line with previous reports on the evoked GBA (Galambos et al., 1981; Tiitinen et al., 1993; Debener et al., 2003; Senkowski et al., 2005; Schadow et al., 2007), time-frequency windows were selected between 25 and 35 Hz and 50–100 ms after auditory stimulus onset. The statistical analysis of evoked GBA included one frontocentral ROI, for which effects of the factor congruency (congruent, incongruent) were examined.

Spectral analyses: total power. TFRs were computed using the multitaper method applied to short sliding time windows, which achieves an optimal concentration of spectral energy over the frequency range of interest (Mitra and Pesaran, 1999). For the high frequencies (20–100 Hz), the data in each time window were multiplied with a set of orthogonal Slepian tapers. The Fourier transforms of the tapered time windows were then calculated, and the resulting power estimates were averaged across tapers. For the multitaper analysis, a fixed time window ($\Delta T = 200$ ms) and a fixed frequency smoothing ($\Delta f = \pm 10$ Hz) was applied, resulting in three tapers being applied to the sliding time window. For the

calculation of total power frequency, transformations were performed before averaging on the single-trial level separately for each frequency. Thus, the resulting total power contains signal components both phase locked and non-phase locked to the stimulus. To reveal the relative signal change of total power in the poststimulus compared with the prestimulus interval, the power in the baseline interval (−300 to −100 ms before stimulus onset) was first subtracted, and the resulting difference was divided by the baseline interval activity as follows: $P(t,f)_{corrected} = 100 \times (P(t,f)_{poststimulus} - P(f)_{baseline})/P(f)_{baseline}$. Total power activity was calculated for the same seven standard ROIs used in the ERP analysis. For statistical analysis of the total power, a cluster-based randomization test proposed by Maris and Oostenveld (2007) was applied. This test controls the type I error rate in experimental designs involving multiple comparisons (across ROIs and time-frequency points) by clustering adjacent ROIs and time-frequency points exhibiting the same effect. First, a dependent sample *t* test was calculated between conditions for all ROI–time-frequency points. The results of the *t* test were then thresholded at an α level of $p = 0.05$. Contiguous ROI–time-frequency points exceeding the threshold were grouped into clusters. Second, each cluster was subjected to a cluster-level test: for a given cluster, the sum of the ROI–time-frequency point *t* values was used in the cluster-level test statistic. The Monte Carlo estimate of the permutation *p* value of the cluster was obtained by comparing the cluster-level test statistic to a randomization-null distribution, assuming no difference between the conditions. This distribution was obtained by 1000 times randomly swapping the conditions in subjects and calculating the maximum cluster-level test statistic. Using 1000 draws, the Monte Carlo *p* value is an accurate estimate of the true *p* value. Grand mean time-frequency representations were computed across all participants to illustrate the relative change of activity to baseline across frequencies. For exploratory purposes, effects of semantic congruency were also examined in lower-frequency responses (<20 Hz; see supplemental material, available at www.jneurosci.org as supplemental material).

Source reconstruction of frequency-specific activity. A linear beamforming approach was applied to reconstruct the cortical sources for the oscillatory GBA (Van Veen et al., 1997; Gross et al., 2001). This source reconstruction technique uses an adaptive spatial filter, which passes activity from one specific location of interest with unit gain and maximally suppresses activity from other locations. Since linear beamforming is based on the calculation of the covariance matrix between single channels over trials, this approach is in particular suitable for the analysis of total power. Recent studies have successfully applied linear beamforming for reconstructing the sources of frequency specific activity in magneto-encephalography (Bauer et al., 2006; Hoogenboom et al., 2006; Siegel et al., 2008) and in EEG data (Schneider et al., 2008b).

For the source reconstructions, a volume conduction model was derived from a Montreal Neurological Institute (http://www.mni.mcgill.ca) template brain, resulting in an anatomically realistic three-shell model. The leadfield matrix was calculated using the boundary element method for each grid point in the brain on a regular 6 mm grid. The source activity at each grid point was estimated by constructing a spatial filter using the leadfield at this point and the cross-spectral density matrix. For each participant and condition, the cross-spectral density matrix was calculated between all 124 scalp EEG channels, and separately for a baseline (−200–0 ms, 60 Hz) and a poststimulus (200–400 ms, 60 Hz) time-frequency interval. The source-level analysis was restricted to the significant time-frequency interval of the scalp-level analysis. Across subjects, a paired *t* test was performed for the estimated GBA to examine differences between prestimulus and poststimulus activity on auditory stimuli as well as differences between the congruent and the incongruent conditions. The *t* values were subsequently transformed to *z*-scores and are reported uncorrected. Results of the source reconstruction analysis are given in Talairach coordinates (Talairach and Tournoux, 1988).

Methods of the unisensory control experiment. Twenty-two participants (5 male; mean age, 23.08 years; age range, 20–29 years) were tested in a unisensory auditory S1-S2 priming paradigm. Two sounds of objects (400 ms duration each) were presented either in a congruent condition or in an incongruent condition with a stimulus onset asynchrony of 1000 ms. In the congruent condition, two semantically congruent stimuli (i.e.,
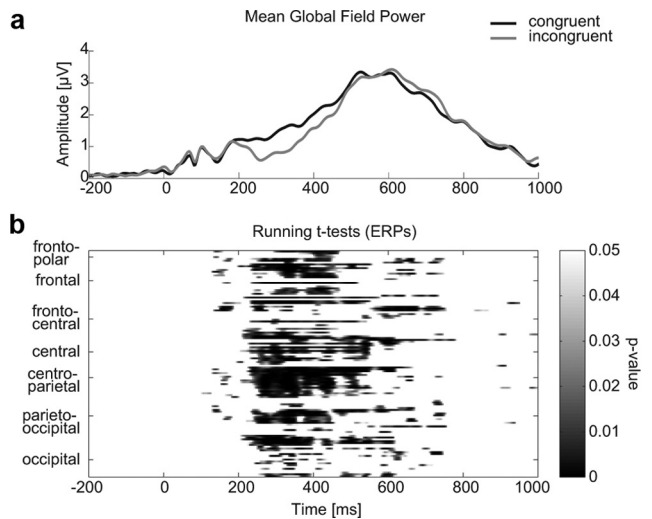
Figure 2. *a*, Group-averaged mean global field power in response to the auditory S2. *b*, The results of point-wise *t* tests across all electrode sites reveal earliest differences between 120 and 170 ms, and a temporally sustained difference after 200 ms after stimulus onset.
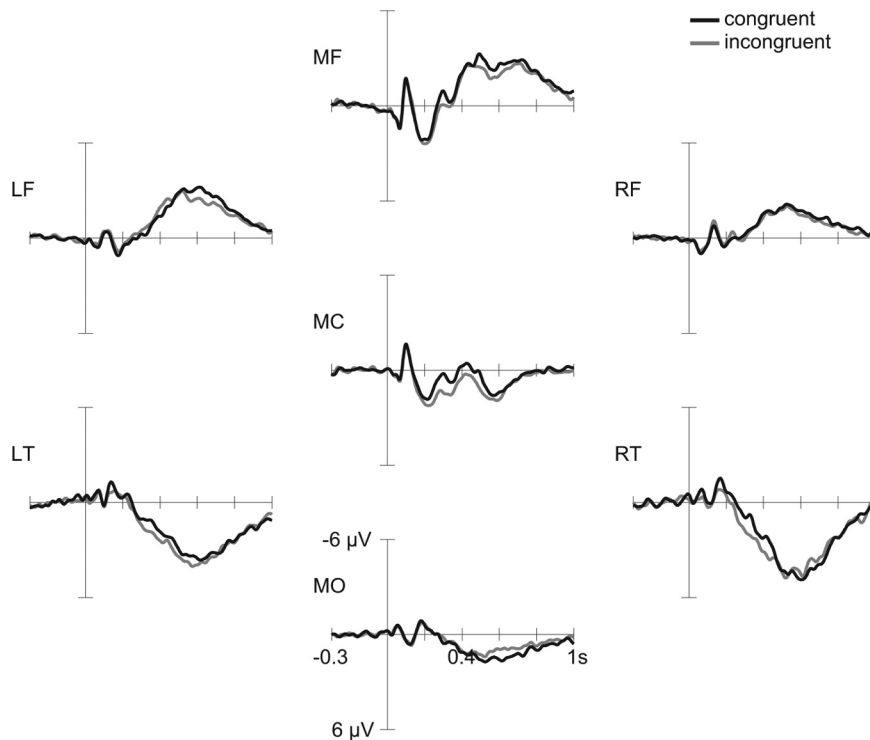

Figure 3. Grand mean ERPs in response to the auditory S2 following semantically congruent and semantically incongruent haptic primes averaged across channels in seven ROIs. L, Left; M, middle; R, right; F, frontal; C, central; O, occipital; T, temporal.

S1 and S2 derived from the same object and physically identical) were presented. In the incongruent condition, two semantically incongruent stimuli were presented (i.e., S1 and S2 were derived from different objects). All auditory stimuli used in this experiment were taken from a standardized stimulus set described previously (Schneider et al., 2008; www.multimost.com). The majority (> 50%) of auditory stimuli overlapped with the auditory stimulus set used in the haptic-to-auditory priming paradigm. Congruent and incongruent stimulus sets were matched by adjusting the mean values in the variables familiarity, identification, and name agreement of each set. The EEG was recorded using 126 channels, and the analysis of total and evoked oscillatory responses was conducted in the same way as done in the haptic-to-auditory priming experiment.

Results

Behavioral data

Analysis of reaction times to the auditory S2 stimuli revealed a cross-modal priming effect as reflected in shorter reaction times for congruent (mean = 452 ms) compared with incongruent stimulus pairs (mean = 486 ms, $t_{(13)} = -5.33$, $p < 0.001$, SD = 23.53, paired sample *t* test). Additionally, error rates differed significantly between the congruent (mean = 5.1%) and the incongruent condition (mean = 8.38%, $t_{(13)} = -4.47$, $p < 0.001$, SD = 2.74, paired sample *t* test). These findings demonstrate a cross-modal priming effect for semantically congruent haptic and auditory inputs.

Event-related potentials

The analysis of GFP revealed amplitude maxima at 60 and 120 ms, followed by a long-lasting potential between 200 and 1000 ms (Fig. 2). Based on the GFP data, three time windows were selected for the further analyses: 60–80, 120–170, and 200–400 ms. To avoid alterations due to auditory stimulus offset and preparatory motor responses, the statistical analysis of the ERPs was restricted to the interval between 0 and 400 ms following sound onset. The ERPs in response to auditory stimuli are shown in Figure 3 for all seven ROIs, revealing a centrally distributed N1 (~100 ms) and a P2 (~200 ms) component.

In the 60–80 ms time window, the ANOVA with factors congruency and region revealed only an effect of the factor region ($F_{(1,13)} = 14.6$, $p < 0.001$), but no significant main effect of congruency ($F_{(1,13)} = 0.66$, $p = 0.43$) and no interaction ($F_{(6,78)} = 1.43$, $p = 0.22$).

In the 120–170 ms time window, the ANOVA with factors congruency and region revealed a significant main effect of region ($F_{(1,13)} = 3.81$, $p < 0.01$) and a significant congruency × region interaction ($F_{(6,78)} = 3.24$, $p < 0.01$). Follow-up ANOVAs were performed to test the effects of semantic congruency at each of seven ROIs separately. A significant effect of congruency was found at the left frontal ROI ($F_{(1,13)} = 6.86$, $p < 0.05$). A significant effect was also observed at the left temporal ROI ($F_{(1,13)} = 4.72$, $p < 0.05$) and a trend toward significance at the right temporal ROI ($F_{(1,13)} = 3.15$, $p = 0.09$), with both temporal ROIs showing a larger negativity for the incongruent condition.

In the 200–400 ms time window, significant main effects of congruency ($F_{(1,13)} = 31.98$, $p < 0.001$) and region ($F_{(1,13)} = 4.34$, $p < 0.001$), and a significant congruency × region interaction ($F_{(6,78)} = 8.56$, $p = 0.001$) were found. In follow-up ANOVAs for the ROIs separately, a significant effect of congruency was found at the middle frontal ROI ($F_{(1,13)} = 8.78$, $p < 0.05$), the middle central ROI ($F_{(1,13)} = 27.93$, $p < 0.001$), the left temporal ROI ($F_{(1,13)} = 5.62$, $p < 0.05$), and the right temporal ROI ($F_{(1,13)} = 26.94$, $p < 0.001$) due to larger negative deflections in the incongruent compared with the congruent condition. The topographical distribution of the ERPs in the two time windows between 120 and 170 ms and 200–400 ms are depicted in Figure 4.

Frequency-specific power changes

Evoked power

Figure 5 shows the time-frequency representation of evoked frontocentral activity following the auditory stimulus. A peak between 50 and 100 ms is present, showing a maximum at frontocentral sensors at a frequency of ~25–35 Hz. The repeated-measures ANOVA for the frontocentral ROI, which was selected ad hoc based on previous reports, revealed enhanced frontocentral GBA for congruent compared with incongruent trials ($F_{(1,13)} = 5.67, p < 0.05$). An exploratory investigation of other ROIs, which have been tested in the above-described analysis of ERPs, did not reveal any other significant effects. Thus, the effect of semantic congruency on the evoked GBA is restricted to frontocentral sensors.

Analysis of the intertrial coherence (Fig. 6) in the respective time-frequency window of the evoked GBA effect revealed significant frontocentral differences between the congruent and the incongruent conditions ($F_{(1,13)} = 5.23, p < 0.05$), showing stronger phase resetting for the congruent compared with the incongruent condition.

Total power

The time-frequency representations of total power showed a clear increase in GBA for a frequency range of 60–70 Hz in the 200–400 ms time window (Fig. 7). A randomization approach correcting for multiple comparisons was applied to access significance between conditions. The randomization routine identified a significant cluster at the frontocentral and middle central ROIs (60–70 Hz, 250–350 ms). Gamma power in this time window was enhanced for the congruent compared with the incongruent condition. For this specific time-frequency window, the source reconstruction using linear beamforming was applied. Analysis of the intertrial coherence (Fig. 6) in the respective time-frequency window (60–70 Hz, 250–350 ms) revealed no significant difference between the congruent and the incongruent condition, indicating that the congruency effect observed in total power is mainly due to a power change in the non-phase-locked part of the GBA. The exploratory analysis of lower-frequency band responses revealed a significantly stronger power suppression (200–450 ms, 8–16 Hz) at the middle central and middle frontal ROIs for congruent compared with incongruent trials (supplemental Fig. 1, available at www.jneurosci.org as supplemental material). In the unisensory auditory control experiment, no significant differences between congruent and incongruent trials were observed in the total GBA (supplemental Fig. 2, available at www.jneurosci.org as supplemental material). A repeated-measures ANOVA comprising the within-subject factor congruency (congruent, incongruent) and the between-subject factor experiment (multisensory, unisensory) revealed a significant interaction between the two fac-
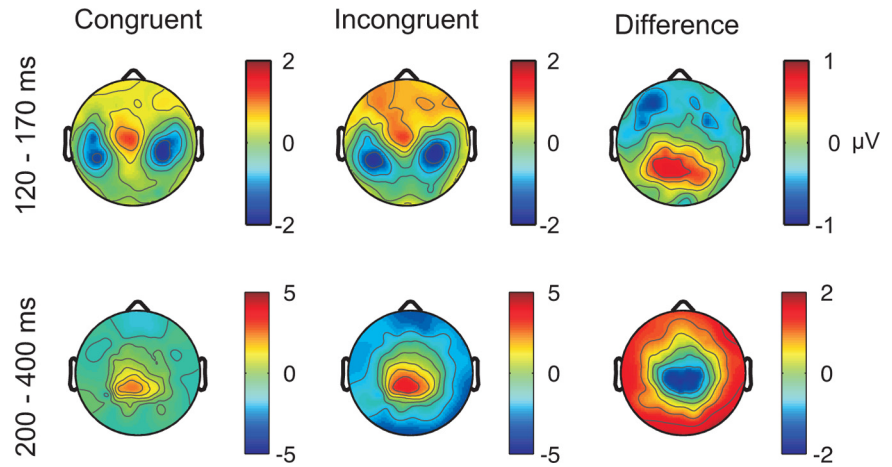

Figure 4. Topographic maps of the congruent, the incongruent, and the difference group-averaged ERPs in the two significant time windows.
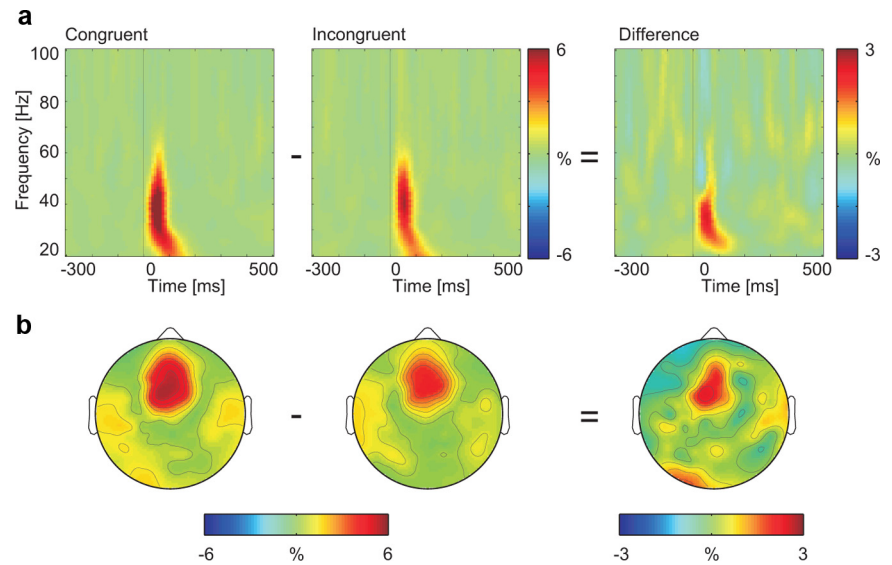

Figure 5. *a*, Group-averaged time-frequency representation of the evoked GBA at the frontocentral ROI in response to the auditory S2 in the semantically congruent and incongruent conditions and the difference between the conditions. *b*, Topographic maps of the group averaged evoked GBA (25–35 Hz) between 50 and 100 ms.
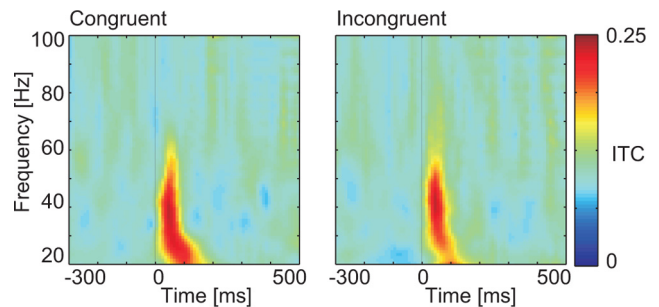

Figure 6. Group averaged ITC in response to the auditory S2 in the semantically congruent (left) and incongruent (right) condition.

tors ($F_{(1,34)} = 7.601, p < 0.01$, partial $\eta^2 = 0.183$; see supplemental material, available at www.jneurosci.org as supplemental material), showing that the effect of congruency on total GBA is specific to the cross-modal priming experiment.
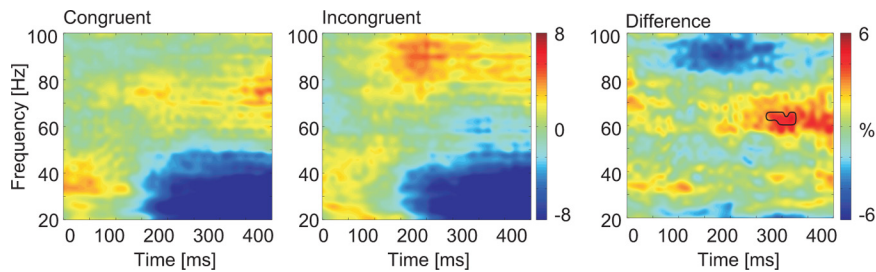
Figure 7. Group-averaged time-frequency representation of the total GBA at the middle central ROI in response to the auditory S2 in the semantically congruent and incongruent conditions as well as the difference between the conditions.
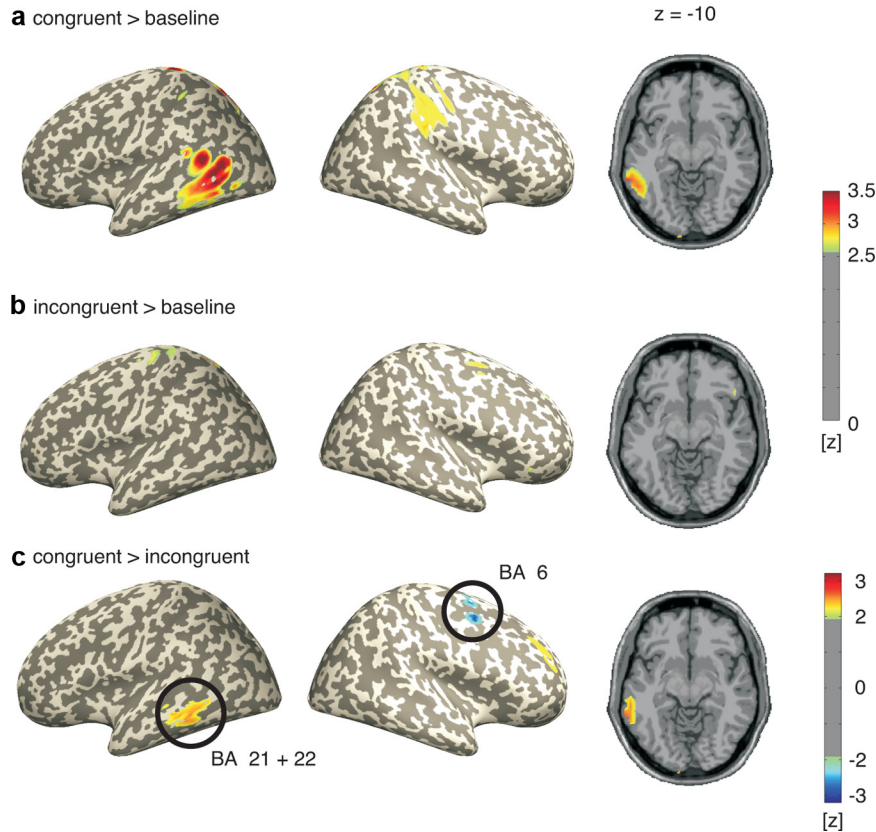

Figure 8. *a*, *b*, Results of source reconstruction using linear beamforming for the total GBA (60 Hz) in the congruent (*a*) and incongruent (*b*) conditions tested against prestimulus baseline activity. *c*, Statistical differences between the congruent and the incongruent condition across subjects; warm colors indicate stronger GBA in the congruent condition; cold colors indicate stronger GBA in the incongruent condition.

Source analysis of frequency-specific activity

The source reconstruction for total GBA to the auditory S2 in congruent trials (Fig. 8*a*) revealed an increase of activity in the right paracentral lobule [Brodmann's area (BA) 6; coordinates in mm, peak: 3/−30/69; $z = 3.99$)] and in the left superior temporal gyrus/superior temporal sulcus (STG/STS) (BA 22; peak: −57/−52/11; $z = 3.63$). In the incongruent condition (Fig. 8*b*), an increase in GBA was found in the right superior frontal gyrus (BA 8; peak: 9/51/43; $z = 3.46$). Importantly, the direct comparison between the two conditions (Fig. 8*c*) revealed significant differences in left MTG (BA 21; peak: −61/ −40/−10; $z = 2.32$) and right middle frontal gyrus (BA 6; peak: 38/5/56; $z = -2.74$). The left middle temporal responses in the gamma band were enhanced in the congruent compared with the incongruent condition; whereas, the right middle frontal gyrus

GBA was stronger in the incongruent compared with the congruent condition.

Discussion

We have characterized the modulation of oscillatory neuronal activity during auditory object categorization influenced by haptic sensory information in a cross-modal priming paradigm. On the behavioral level, cross-modal priming was reflected in shorter reaction times and lower error rates for congruent compared with incongruent haptic–auditory stimulus pairs. On the physiological level, we found a significant GBA modulation in areas of multisensory convergence to semantically congruent compared with incongruent inputs, which extends recent findings demonstrating an increase of GBA associated with visual–auditory cross-modal semantic priming (Schneider et al., 2008b). In addition, short- and long-latency effects of semantic congruency were also identified in the ERPs, possibly reflecting multiple stages of cross-modal semantic matching and contextual integration.

Event-related potentials

In the present experiment, the earliest congruency effect in the ERPs associated with semantic matching between the haptic and the auditory information was observed between 120 and 170 ms. This effect is likely to reflect an early cross-modal matching process, as signs of multisensory integration have been reported with similar latency in recent studies using naturalistic stimuli (Senkowski et al., 2007a; Stekelenburg and Vroomen, 2007). This suggests that cross-modal matching can occur rapidly in the auditory processing stream.

The late effects in the 200 to 400 ms time window are in close agreement with our previous findings of a negative-going difference wave (250–350 ms) with a central posterior distribution for incongruent compared with congruent stimulation (Schneider et al., 2008b). We interpret this effect to reflect contextual integration processes triggered by deviations from the semantic context. In accordance with the present findings, ERP effects starting 200 ms after stimulus onset were also observed in cross-modal semantic priming experiments using visually presented words and environmental sounds (Orgs et al., 2006) as well as acoustically and visually presented words (Holcomb and Anderson, 1993; Holcomb et al., 2005), interpreted as N400. The N400 component is hypothesized to reflect "contextual integration" (Brown and Hagoort, 1993; Kutas and Federmeier, 2000) as it is increased in amplitude in case target stimuli are semantically unrelated to the local context. Importantly, in the present study it becomes evident that semantically incongruent inputs to the haptic and the auditory modality are reflected in a long-lasting N400 effect. We conclude that this ERP

effect reflects higher cognitive operations, such as semantic context integration, which are probably elicited independent of modality.

Frequency-specific power changes

Two distinct effects of cross-modal semantic priming were observed in GBA. First, frontocentral evoked GBA was enhanced for semantically congruent compared with incongruent haptic–auditory stimulation at a short latency (50–100 ms; 25–35 Hz). Second, total power activity in the gamma band (60–70 Hz) was modulated by semantic congruency at a later time window (250–350 ms).

Short-latency evoked power effects
Simple auditory stimuli have been shown to generate an early evoked response (15–150 ms) in the low gamma band (20–40 Hz) in the human temporal cortex, as observed by intracranial recordings (Edwards et al., 2005). Experiments using multisensory, audiovisual stimulation have revealed that the evoked gamma-band response in the EEG is modulated by multiple factors including selective attention (Senkowski et al., 2005) and relative stimulus onset timing (Senkowski et al., 2007b). The early evoked enhancement of GBA to semantically congruent stimuli in our study may reflect a match between the expected and the actual stimulus. This interpretation accords well with the memory match and utilization model (Herrmann et al., 2004), which posits that early evoked GBA is increased during a match between stimulus-related bottom-up information and working-memory top-down information.

The findings of early evoked GBA modulations seem to somewhat contradict our recent findings on implicit visual-to-auditory object priming (Schneider et al., 2008b) and explicit visual-to-auditory object matching (Senkowski et al., 2009), revealing no effects on early evoked GBA. The differences in results may be related to the following distinct features of the present study: (1) participants were explicitly familiarized with the stimuli in an extensive learning phase; (2) the number of repetitions of each individual stimulus was higher potentially producing more constrained expectations; and (3) haptic in contrast to visual prime stimuli have been used. Future studies will have to elucidate which factors produce modulation of early evoked GBA in cross-modal processing. Moreover, since the comparison of priming effects on evoked GBA between the multisensory and the unisensory priming experiments did not reveal significant differences, it remains to be investigated under which conditions those effects may be specific to unisensory or multisensory processing.

Source estimations of magnetoencephalography recordings have suggested that the early evoked gamma response to auditory stimuli is partly generated in primary auditory areas (Pantev et al., 1991). The frontocentral topography of the early evoked GBA in the EEG is compatible with this result, as it could be explained by two tangential dipoles in or near superior temporal cortex. Thus, the modulation of evoked GBA may reflect a match between the expected and the actual stimulus in early sensory areas, which also accords with the observation made in a study on monkeys that somatosensory inputs can influence oscillatory gamma responses in primary auditory cortex (Lakatos et al., 2007). This study suggests that the cross-modal priming effect observed here might be linked to a phase reset of oscillatory activity in the gamma band. Comparably, a recent study reported an influence of attention on oscillatory phase reset in primary auditory and visual cortices (Lakatos et al., 2009).

Long-latency total power effects
Total power GBA was enhanced in response to semantically congruent cross-modal information between the haptic and the auditory systems. This activity may reflect semantic matching processes that are particularly associated with cross-modal priming. The absence of such effects in the unisensory auditory priming experiment supports this assumption. Furthermore, in our previous investigations enhanced total power GBA was indicative of implicit (Schneider et al., 2008b) and explicit (Senkowski et al., 2009) semantic matching between the visual and the auditory modalities.

Long-latency (> 200 ms) activity in the gamma band has been frequently associated with memory processes (Gruber and Müller, 2005). According to the match and utilization model (Herrmann et al., 2004), longer-latency non-phase-locked GBA may be a signature of a utilization process as, for example, the selection of an appropriate behavioral response. Recently, coherent oscillatory activity has been proposed as a functional mechanism for multisensory processing (Maier et al., 2008; Senkowski et al., 2008). According to this view, cross-modal matching processes could be reflected in enhanced oscillatory neural activity, as observed in response to semantically congruent inputs through different sensory modalities.

The results of the source analysis revealed that oscillatory activity in left STG/STS and middle temporal gyrus (BA 21/22) is associated with cross-modal semantic priming. Interestingly, in our previous investigation neighboring and partly overlapping areas (BA 21/20) were likewise activated by semantically congruent stimulation in visual-to-auditory priming (Schneider et al., 2008b). In accordance with these findings, the STS was suggested to serve as a multisensory convergence area for auditory and somatosensory inputs (Foxe et al., 2002). Most importantly, the lateral temporal cortex, including STG/STS and MTG, has been repeatedly identified as an area of multisensory convergence during semantically congruent stimulation (Doehrmann and Naumer, 2008). Together, our data provide further evidence that the lateral temporal cortex plays a functional role in multisensory object processing and that oscillatory responses in the gamma band are linked to cross-modal semantic matching in these structures.

Additionally, enhanced GABA was found in the present study for semantically incongruent haptic–auditory stimulation in medial frontal cortex (MFC). During processing of incongruent auditory–visual stimuli, stronger activations were observed in MFC and the anterior cingulate cortex (Taylor et al., 2006; Noppeney et al., 2008). These areas have been implicated in the detection of conflicts between simultaneously active representations of stimuli as well as error processing and cognitive control (Botvinick et al., 2004; Ridderinkhof et al., 2004).

Finally, an exploratory analysis of lower-frequency band responses revealed a stronger poststimulus power decrease in the α and lower β band range (8–16 Hz) for congruent compared with incongruent trials. Such decreases in low-frequency bands accompanying power enhancements in the gamma band have been observed in numerous studies on sensory processing and on attentional modulation of stimulus saliency (e.g., Bauer et al., 2006; Siegel et al., 2007, 2008; Engel and Fries, 2010), suggesting that the same pattern of power shifts from the low- to high-frequency band activity also occurs in cross-modal semantic priming.

Conclusion

Our study provides evidence that haptic object exploration facilitates the recognition of meaningful auditory stimuli in haptic-to-auditory priming of objects. The physiological data suggest

that semantically matching haptic and auditory signals are integrated in a network of unisensory and multisensory regions. The matching operations performed by this network seem to be reflected in the spatiotemporal dynamics of processing at multiple stages. The observed enhancement of the early evoked GBA (50–100 ms) possibly reflects the modulation of auditory stimulus processing by haptic inputs at early stages of information processing. Cross-modal matching processes were reflected in changes of total GBA (250–350 ms). The effect in total GBA was localized to distinct regions located in the lateral temporal lobe, which are known to be involved in multisensory semantic processing. The findings on GBA were complemented by cross-modal priming effects in the ERPs. Thus, our data suggest the operation of a cortical network in which coherent oscillatory signals may serve as a mechanism for the integration of sensory inputs across modalities.

References

Amedi A, Malach R, Hendler T, Peled S, Zohary E (2001) Visuo-haptic object-related activation in the ventral visual pathway. Nat Neurosci 4:324–330.

Bauer M, Oostenveld R, Peeters M, Fries P (2006) Tactile spatial attention enhances gamma-band activity in somatosensory cortex and reduces low-frequency activity in parieto-occipital areas. J Neurosci 26:490–501.

Botvinick MM, Cohen JD, Carter CS (2004) Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn Sci 8:539–546.

Brown C, Hagoort P (1993) The processing nature of the N400: evidence from masked priming. J Cogn Neurosci 5:33–44.

Debener S, Herrmann CS, Kranczioch C, Gembris D, Engel AK (2003) Top-down attentional processing enhances auditory evoked gamma band activity. Neuroreport 14:683–686.

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.

Doehrmann O, Naumer MJ (2008) Semantics and the multisensory brain: how meaning modulates processes of audio-visual integration. Brain Res 1242:136–150.

Easton RD, Srinivas K, Greene AJ (1997) Do vision and haptics share common representations? Implicit and explicit memory within and between modalities. J Exp Psychol Learn Mem Cogn 23:153–163.

Edwards E, Soltani M, Deouell LY, Berger MS, Knight RT (2005) High gamma activity in response to deviant auditory stimuli recorded directly from human cortex. J Neurophysiol 94:4269–4280.

Engel AK, Fries P (2010) Beta-band oscillations - signalling the status quo? Curr Opin Neurobiol 20:156–165.

Engel AK, König P, Kreiter AK, Schillen TB, Singer W (1992) Temporal coding in the visual cortex: new vistas on integration in the nervous system. Trends Neurosci 15:218–226.

Engel AK, Fries P, Singer W (2001) Dynamic predictions: oscillations and synchrony in top-down processing. Nat Rev Neurosci 2:704–716.

Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, Ritter W, Murray MM (2002) Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. J Neurophysiol 88:540–543.

Fu KM, Johnston TA, Shah AS, Arnold L, Smiley J, Hackett TA, Garraghty PE, Schroeder CE (2003) Auditory cortical neurons respond to somatosensory stimulation. J Neurosci 23:7510–7515.

Galambos R, Makeig S, Talmachoff PJ (1981) A 40-Hz auditory potential recorded from the human scalp. Proc Natl Acad Sci U S A 78:2643–2647.

Gross J, Kujala J, Hamalainen M, Timmermann L, Schnitzler A, Salmelin R (2001) Dynamic imaging of coherent sources: studying neural interactions in the human brain. Proc Natl Acad Sci U S A 98:694–699.

Gruber T, Müller MM (2005) Oscillatory brain activity dissociates between associative stimulus content in a repetition priming task in the human EEG. Cereb Cortex 15:109–116.

Guthrie D, Buchwald JS (1991) Significance testing of difference potentials. Psychophysiology 28:240–244.

Herrmann CS, Munk MH, Engel AK (2004) Cognitive functions of gamma-band activity: memory match and utilization. Trends Cogn Sci 8:347–355.

Holcomb PJ, Anderson JE (1993) Cross-modal semantic priming: a time course analysis using event-related brain potentials. Lang Cogn Process 8:379–411.

Holcomb PJ, Anderson J, Grainger J (2005) An electrophysiological study of cross-modal repetition priming. Psychophysiology 42:493–507.

Hoogenboom N, Schoffelen JM, Oostenveld R, Parkes LM, Fries P (2006) Localizing human visual gamma-band activity in frequency, time and space. Neuroimage 29:764–773.

James TW, Humphrey GK, Gati JS, Servos P, Menon RS, Goodale MA (2002) Haptic study of three-dimensional objects activates extrastriate visual areas. Neuropsychologia 40:1706–1714.

Kaiser J, Hertrich I, Ackermann H, Mathiak K, Lutzenberger W (2005) Hearing lips: gamma-band activity during audiovisual speech perception. Cereb Cortex 15:646–653.

Kayser C, Petkov CI, Augath M, Logothetis NK (2005) Integration of touch and sound in auditory cortex. Neuron 48:373–384.

Kayser C, Petkov CI, Logothetis NK (2008) Visual modulation of neurons in auditory cortex. Cereb Cortex 18:1560–1574.

Kayser C, Petkov CI, Logothetis NK (2009) Multisensory interactions in primate auditory cortex: fMRI and electrophysiology. Hear Res 258:80–88.

Kutas M, Federmeier KD (2000) Electrophysiology reveals semantic memory use in language comprehension. Trends Cogn Sci 4:463–470.

Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE (2007) Neuronal oscillations and multisensory interaction in primary auditory cortex. Neuron 53:279–292.

Lakatos P, O'Connell MN, Barczak A, Mills A, Javitt DC, Schroeder CE (2009) The leading sense: supramodal control of neurophysiological context by attention. Neuron 64:419–430.

Lederman SJ, Klatzky RL (1987) Hand movements: a window into haptic object recognition. Cogn Psychol 19:342–368.

Lederman SJ, Klatzky RL (1993) Extracting object properties through haptic exploration. Acta Psychol (Amst) 84:29–40.

Lehmann D, Skrandies W (1980) Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr Clin Neurophysiol 48:609–621.

Maier JX, Chandrasekaran C, Ghazanfar AA (2008) Integration of bimodal looming signals through neuronal coherence in the temporal lobe. Curr Biol 18:963–968.

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164:177–190.

Mitra PP, Pesaran B (1999) Analysis of dynamic brain imaging data. Biophys J 76:691–708.

Noppeney U, Josephs O, Hocking J, Price CJ, Friston KJ (2008) The effect of prior visual information on recognition of speech and sounds. Cereb Cortex 18:598–609.

Orgs G, Lange K, Dombrowski JH, Heil M (2006) Conceptual priming for environmental sounds and words: an ERP study. Brain Cogn 62:267–272.

Pantev C, Makeig S, Hoke M, Galambos R, Hampson S, Gallen C (1991) Human auditory evoked gamma-band magnetic fields. Proc Natl Acad Sci U S A 88:8996–9000.

Peltier S, Stilla R, Mariola E, LaConte S, Hu X, Sathian K (2007) Activity and effective connectivity of parietal and occipital cortical regions during haptic shape perception. Neuropsychologia 45:476–483.

Reales JM, Ballesteros S (1999) Implicit and explicit memory for visual and haptic objects: cross-modal priming depends on structural descriptions. J Exp Psychol Learn Mem Cogn 25:644–663.

Renier LA, Anurova I, De Volder AG, Carlson S, VanMeter J, Rauschecker JP (2009) Multisensory integration of sounds and vibrotactile stimuli in processing streams for "what" and "where." J Neurosci 29:10950–10960.

Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S (2004) The role of the medial frontal cortex in cognitive control. Science 306:443–447.

Schadow J, Lenz D, Thaerig S, Busch NA, Fründ I, Herrmann CS (2007) Stimulus intensity affects early sensory processing: sound intensity modulates auditory evoked gamma-band activity in human EEG. Int J Psychophysiol 65:152–161.

Schneider TR, Engel AK, Debener S (2008a) Multisensory identification of natural objects in a two-way crossmodal priming paradigm. Exp Psychol 55:121–132.

Schneider TR, Debener S, Oostenveld R, Engel AK (2008b) Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. Neuroimage 42:1244–1254.

Schroeder CE, Lindsley RW, Specht C, Marcovici A, Smiley JF, Javitt DC (2001) Somatosensory input to auditory association cortex in the macaque monkey. J Neurophysiol 85:1322–1327.

Senkowski D, Talsma D, Herrmann CS, Woldorff MG (2005) Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. Exp Brain Res 166:411–426.

Senkowski D, Saint-Amour D, Kelly SP, Foxe JJ (2007a) Multisensory processing of naturalistic objects in motion: a high-density electrical mapping and source estimation study. Neuroimage 36:877–888.

Senkowski D, Talsma D, Grigutsch M, Herrmann CS, Woldorff MG (2007b) Good times for multisensory integration: Effects of the precision of temporal synchrony as revealed by gamma-band oscillations. Neuropsychologia 45:561–571.

Senkowski D, Schneider TR, Foxe JJ, Engel AK (2008) Crossmodal binding through neural coherence: implications for multisensory processing. Trends Neurosci 31:401–409.

Senkowski D, Schneider TR, Tandler F, Engel AK (2009) Gamma-band activity reflects multisensory matching in working memory. Exp Brain Res 198:363–372.

Siegel M, Donner TH, Oostenveld R, Fries P, Engel AK (2008) Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. Neuron 60:709–719.

Singer W, Gray CM (1995) Visual feature integration and the temporal correlation hypothesis. Annu Rev Neurosci 18:555–586.

Stekelenburg JJ, Vroomen J (2007) Neural correlates of multisensory integration of ecologically valid audiovisual events. J Cogn Neurosci 19:1964–1973.

Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. New York: Thieme Medical Publishers.

Taylor KI, Moss HE, Stamatakis EA, Tyler LK (2006) Binding crossmodal object features in perirhinal cortex. Proc Natl Acad Sci U S A 103:8239–8244.

Tiitinen H, Sinkkonen J, Reinikainen K, Alho K, Lavikainen J, Näätänen R (1993) Selective attention enhances the auditory 40-Hz transient response in humans. Nature 364:59–60.

Van Veen BD, van Drongelen W, Yuchtman M, Suzuki A (1997) Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans Biomed Eng 44:867–880.