NeuroImage 42 (2008) 1244–1254

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming

Till R. Schneider [a,*], Stefan Debener [b], Robert Oostenveld [c], Andreas K. Engel [a]

[a] Department of Neurophysiology and Pathophysiology, Center of Experimental Medicine, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany
[b] MRC Institute of Hearing Research Southampton, Southampton Hants, Royal South Hants Hospital, SO14 OYG Southampton, United Kingdom
[c] F.C. Donders Centre for Cognitive Neuroimaging, University Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 25 October 2007
Revised 8 May 2008
Accepted 11 May 2008
Available online 10 July 2008

ABSTRACT

An important step in perceptual processing is the integration of information from different sensory modalities into a coherent percept. It has been suggested that such crossmodal binding might be achieved by transient synchronization of neurons from different modalities in the gamma-frequency range (>30 Hz). Here we employed a crossmodal priming paradigm, modulating the semantic congruency between visual–auditory natural object stimulus pairs, during the recording of the high density electroencephalogram (EEG). Subjects performed a semantic categorization task. Analysis of the behavioral data showed a crossmodal priming effect (facilitated auditory object recognition) in response to semantically congruent stimuli. Differences in event-related potentials (ERP) were found between 250 and 350 ms, which were localized to left middle temporal gyrus (BA 21) using a distributed linear source model. Early gamma-band activity (40–50 Hz) was increased between 120 ms and 180 ms following auditory stimulus onset for semantically congruent stimulus pairs. Source reconstruction for this gamma-band response revealed a maximal increase in left middle temporal gyrus (BA 21), an area known to be related to the processing of both complex auditory stimuli and multisensory processing. The data support the hypothesis that oscillatory activity in the gamma-band reflects crossmodal semantic-matching processes in multisensory convergence sites.

Introduction

Seeing and hearing an object do not necessarily occur at the same moment in a natural environment. Information from one sensory modality can precede information from another by some hundreds of milliseconds. Walking along a street, for example, we may first see or first hear an approaching car; yet we perceive the car as one coherent object. Moreover, the auditory system may be primed by the visual input, and auditory object processing may thus be facilitated. Transient synchronization of neurons has been hypothesized to provide a solution to the problem of integration in distributed networks (von der Malsburg and Schneider, 1986; Singer and Gray, 1995; Engel et al., 2001; Varela et al., 2001). In particular, oscillatory activity in the gamma-band (>30 Hz) is thought to play an important role for the integration of object features and matching of object specific information with memory contents (Tallon-Baudry and Bertrand, 1999; Herrmann et al., 2004). Various studies support the hypothesis of temporal binding by coherent neuronal activity *within* one sensory modality (Gray et al., 1989; Engel et al., 1991a,b). The question of whether temporal synchronization of neural activity may also serve as a mechanism for integration of object features *across* modalities is currently a matter of debate (Senkowski et al., in press). Only a few studies have investigated gamma-band activity in the context of auditory–visual integration (Senkowski et al., 2005, 2007b; Widmann et al., 2007; Yuval-Greenberg and Deouell, 2007).

In order to study the influence of vision on auditory object identification and the possible involvement of temporal binding mechanisms, we employed a crossmodal priming paradigm. We presented visual primes prior to auditory targets and modulated the semantic congruency between the two modalities. Crossmodal semantic matching serves as one important factor in multisensory integration alongside to temporal or spatial correspondence (Laurienti et al., 2004). In a behavioral study we recently observed that prior exposure to semantically congruent visual stimuli facilitated identification of auditory objects, as expressed in shorter reaction times and reduced error rates (Schneider et al., 2008). In the present study, we analyzed the effects of crossmodal semantic priming on both ERPs and on gamma-band responses induced by the task-relevant second stimulus.

* Corresponding author. Fax: +49 40 428037752.
E-mail address: t.schneider@uke.de (T.R. Schneider).

1053-8119/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2008.05.033

Reduced neuronal activity is often reported following repeated processing of the same or similar stimuli. This repetition suppression effect has been observed both at the cellular level and in functional magnetic resonance imaging (fMRI) (Buckner et al., 1998; Wiggs and Martin, 1998; Grill-Spector et al., 2006). Similarly, EEG recordings show that repeated presentation of familiar images of visual objects is accompanied by a reduction in total gamma-band activity (Gruber and Müller, 2002, 2005). Priming effects on ERPs have been observed in studies using semantically related and unrelated stimuli, showing differences in late (>200 ms) components (Van Petten and Rheinfelder, 1995). These findings raise the question of whether similar effects on oscillatory responses and ERP components can also be observed following repeated presentation of the same object in *different* modalities. If semantic object information is stored in a common, amodal memory system, a suppression effect in the gamma-band may be expected following semantically congruent crossmodal stimulation. On the other hand, if coherent oscillatory activity in the gamma-band is relevant for integration or matching of information across modalities, then semantically congruent as compared to incongruent visual–auditory stimulation should enhance gamma-band activity. In addition, late ERP effects should occur that have already been observed in studies of conceptual priming using written words and sounds of objects (Orgs et al., 2006).

Methods

Participants

Twenty-six healthy volunteers participated in the experiment (5 male, mean age 23.08, range 20–29) and received monetary compensation for their participation. All participants were native German speakers, had normal hearing (hearing loss <30 dB) and normal or corrected to normal vision (visus>0.9), and reported no history of neurological or psychiatric illness. The study was conducted in accordance with the Declaration of Helsinki and informed consent was obtained from all participants prior to the recordings. Data from four subjects were discarded due to technical malfunction.

Stimuli

Visual stimuli were taken from a pool of 320 color photographs of natural objects, previously rated in a norming study (Schneider et al., 2008). The stimuli were presented centrally for 400 ms, subtending a visual angle of 8.5° vertically and 9° horizontally. Auditory stimuli were taken from a pool of 270 environmental sounds of natural objects from the same norming study. Physical parameters of both auditory and visual stimuli are well suited to elicit sufficient gamma-band responses according to recent findings (Busch et al., 2004; Schadow et al., 2007). The intensities of the sounds were adjusted by equalizing the root mean square power across all sound files. To avoid on- and offset clicking transients the sound files were windowed with a linear 10 ms rise and fall time. Sounds (22 kHz, 16-bit, mono) were played for 400 ms via Eartone foam-protected air tube earphones (AeroCompany, Indianapolis, IN) at 70dB SPL. All stimuli were presented using Presentation 0.80 (Neurobehavioral Systems, San Francisco, CA).

Experimental paradigm

As illustrated in Fig. 1, each trial consisted of a visual prime (S1) and an auditory target stimulus (S2) presented successively with a stimulus onset asynchrony (SOA) of 1400 ms. The auditory–visual stimulus pairs were either semantically congruent (50%) or semantically incongruent (50%), representing conceptually the same or different objects. The congruent (n = 85) and the incongruent stimulus sets (n = 85) were balanced for familiarity, naming accuracy, and correct object categorization according to the previously obtained stimulus norms (Schneider et al., 2008).

None of the auditory stimuli was used in both the congruent and the incongruent stimulus set. We verified that, on average, there was no difference between the time–frequency profiles of the sounds in the two stimulus sets. To this end, we performed a time–frequency analysis (100 Hz–10 kHz) on the individual sound files with a frequency resolution of 100 Hz and a time resolution of 5 ms. Subsequently, we performed a t-test for differences of the average power values in each time–frequency bin. Significant differences ($p<0.05$) were observed in 389 of 8800 bins, i.e. in less than 5% of the bins, which would be expected by chance.

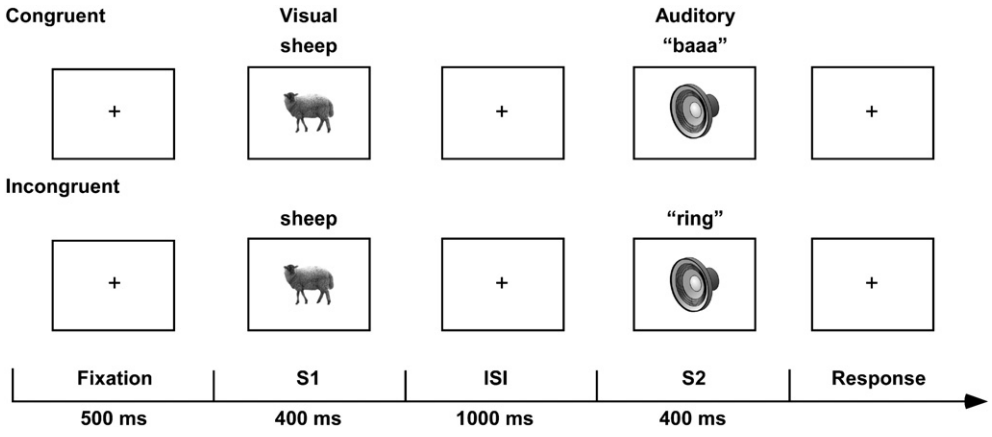

Fig. 1. Illustration of the experimental design. Semantically congruent and incongruent stimuli were randomly presented, each in 50% of the trials. Visual stimuli served as primes (S1) and auditory stimuli served as targets (S2).

Stimuli in both conditions were presented in a pseudo-randomized order, such that the same object-category was never presented in two consecutive trials. The subjects' task was to decide as quickly and as accurately as possible whether the object depicted by the task-relevant stimulus fitted into a shoebox or not. Thus, the task combined the need for object identification with a speeded behavioral response, which is a prerequisite for reaction time analysis. The same task was successfully used in previous priming studies (e.g. Dobbins et al., 2004). Responses had to be given by pressing one of two buttons with the left or the right thumb, counterbalanced across subjects. If no response was given within a time window of 2000 ms, a message was presented reminding the subjects to respond faster in subsequent trials. Violation of the time limit occurred in less than 2.5% of the presented trials. For analysis of the behavioral data according to signal detection theory (Green and Swets, 1966), hit and false alarm rates were computed as follows: correct 'yes' answers to the question 'Does the object fit into a shoebox?' were counted as *hits* and incorrect 'yes' answers to this question were counted as *false alarms*.

EEG recordings

Continuous EEG data were collected from 126 scalp sites using sintered Ag/AgCl ring electrodes mounted on an elastic cap (Falk Minow Services, Herrsching, Germany). During recordings, the nose tip was used as reference, but following artifact rejection the data were re-referenced to average reference (Lehmann and Skrandies, 1980). Two additional electrodes were positioned below the eyes to record the electrooculogram. The data were recorded with an analog passband of 0.016–250 Hz and digitized at a sampling rate of 1000 Hz using BrainAmp amplifiers (BrainProducts, Munich, Germany). Electrode impedances were kept below 20kΩ. Analysis of the data was performed using Matlab 7.01 (MathWorks, Natick, MA) and EEGLAB 4.51 (Delorme and Makeig, 2004), a freely available open source toolbox for EEG data

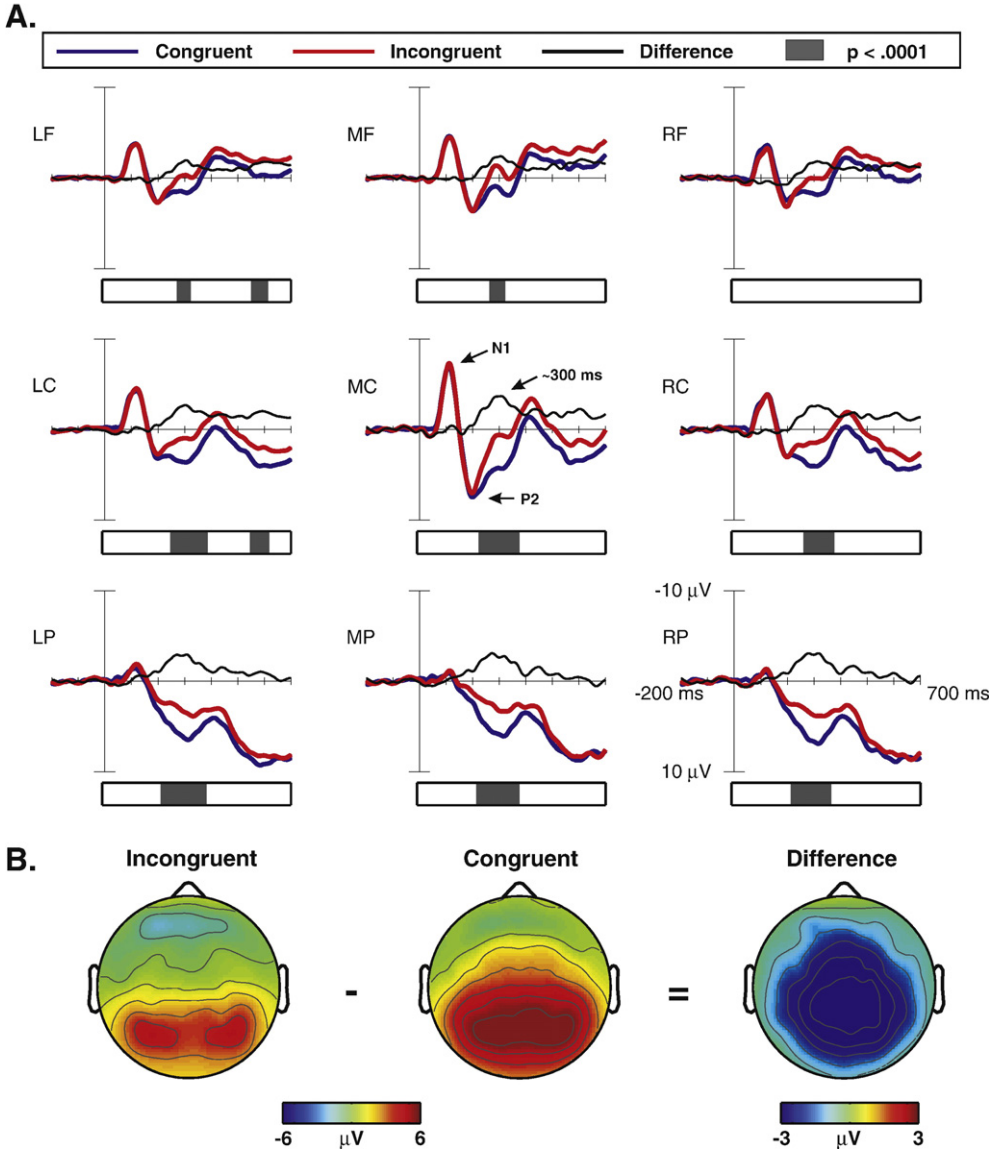

Fig. 2. (A) ERPs to auditory targets (S2) following semantically congruent (blue trace) and semantically incongruent primes (red trace) are shown for the nine regions of interest. Running *t*-tests indicate significant differences between the conditions beginning at approximately 250 ms. (B) The topographic maps of the incongruent, the congruent and the difference ERPs at 250 ms show a central–posterior distribution.

analysis (Swartz Center for Computational Neuroscience, La Jolla, CA; http://www.sccn.ucsd.edu/eeglab). For all further analyses the data were band-pass filtered (0.3–110 Hz) and downsampled to 250 Hz. The data was sampled during the recording at a sampling rate that was higher than actually needed for the analysis. To reduce computational time and memory demands, the data was downsampled prior to the analysis. A two-step procedure was performed to remove artifacts, as used previously (e.g., Debener et al., 2005; Hine and Debener, 2007). First, epochs containing non-stereotyped artifacts (e.g., cable movement, swallowing) were removed. Then, extended infomax independent component analysis (ICA) was applied, using a weight change $<10^{-7}$ as stop criterion. Independent components representing artifacts such as eyeblinks, horizontal eye movements, or electrocardiographic activity were removed from the EEG data by back-projecting all but these components. Application of this procedure ensured that 74.8% of all recorded trials (congruent mean: 75.5%, range: 32% to 97%; incongruent mean: 74%, range: 32% to 97%) could be retained.

Event-related potential analysis

Epochs for the analysis of ERPs were derived starting 200 ms before target onset and lasting for 700 ms. The interval from − 200 ms to stimulus onset served as baseline. A 30 Hz lowpass filter was applied for the ERP analysis. As expected, ERPs in response to S1 did not show any significant differences and we did not further analyze them. Statistical analysis of the ERPs in response to S2 was carried out for nine regions of interest (ROI), each comprising the averaged signal of six adjacent electrodes. Regions were defined as left frontal (LF), middle frontal (MF), right frontal (RF), left central (LC), middle central (MC), right central (RC), left posterior (LP), middle posterior (MP), and right posterior (RP). Statistical analysis was performed in two ways. Firstly, to obtain information about the temporal evolution of significant differences, running *t*-tests were performed testing differences between responses in the congruent and incongruent conditions at each ERP sampling point. An interval was considered to differ significantly between the conditions if at least 11 consecutive data points reached a *p*-value of 0.0001 (Guthrie and Buchwald, 1991). Secondly, a repeated measures ANOVA was performed with the within-subject factors Congruency (congruent, incongruent) and Region of Interest (9 regions) for a response interval from 250 ms to 350 ms. The difference of the topographical distribution between conditions was analyzed computing the global dissimilarity index of the ERP (Lehmann and Skrandies, 1980) using the Cartool software by Denis Brunet (http://brainmapping.unige.ch/Cartool.htm). First, values at each electrode were scaled to unitary strength by dividing with the instantaneous global field power. Subsequently, the square root of the mean of the squared differences at each electrode was computed. For testing the statistical difference between the conditions a bootstrapping test was applied in which the dissimilarity index was calculated for the original data and for

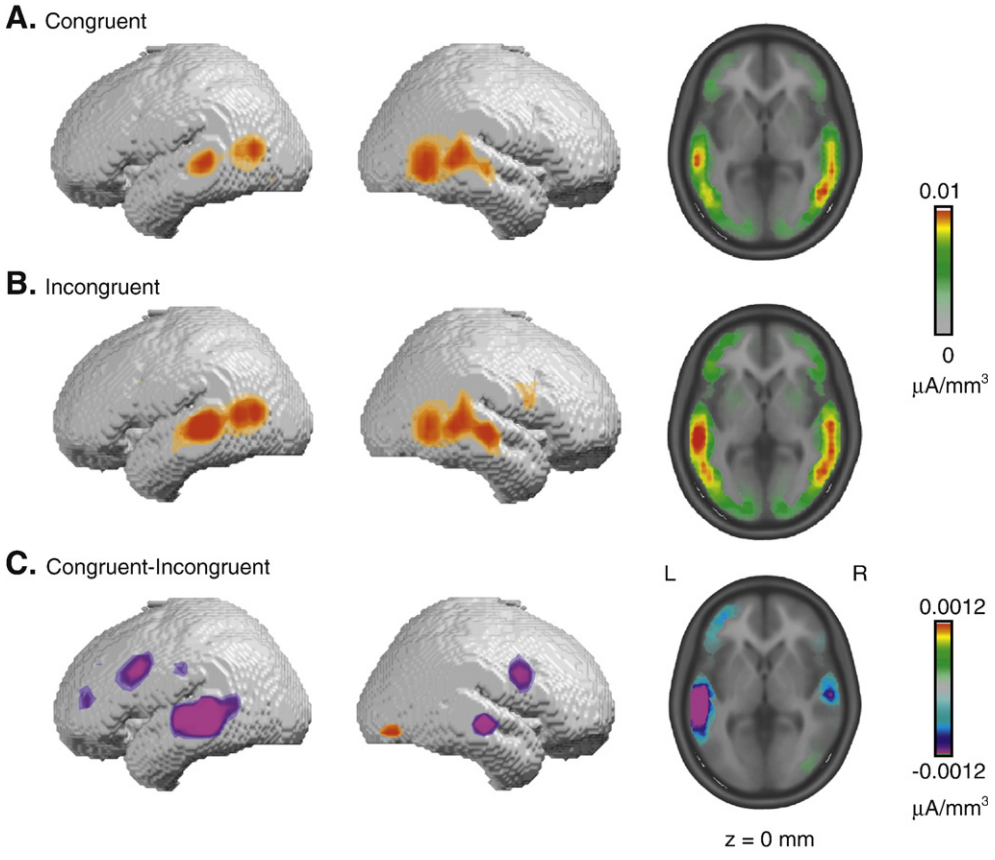

Fig. 3. Source reconstruction of the ERPs in the 250–350 ms time window. The results are plotted on the left and right hemisphere of the MNI brain (left and middle column) and, additionally, on axial slices at $z=0$ (right column). The top two rows (A and B) illustrate the group-averaged source estimates in the semantically congruent and incongruent conditions. The bottom row (C) illustrates the group-averaged source level difference between both conditions.

5000 random reassignments to either of the two experimental conditions (for a similar approach see Murray et al. (2004)).

Source estimations of event-related potentials

Source estimates for the ERP (250–350 ms) were calculated using Local Autoregressive Averages (LAURA) as a distributed linear inverse solution (Grave de Peralta Menendez et al., 2001, 2004) by means of the Cartool software. LAURA is capable of dealing with multiple simultaneously active sources at a priori unknown locations. Technical details of LAURA are described by Grave de Peralta Menendez et al. (2001, 2004). In short, sources were fitted to a realistic head model with a solution space of 4024 nodes selected from a regular 6×6×6 mm grid, restricted to the grey matter of the Montreal Neurological Institute (MNI; http://www.mni.mcgill.ca) template brain. All source localizations are given in Talairach coordinates (Talairach and Tournoux, 1988).

Spectral analysis of gamma-band activity

Spectral changes in oscillatory activity were analyzed using a wavelet transform, which provides a good compromise between time and frequency resolution (Tallon-Baudry and Bertrand, 1999). Time–frequency analysis was performed for each channel by convolving the data with a complex Morlet wavelet $w(t,f_0)$ having a Gaussian shape in time (σ_t) and in frequency (σ_f) around the center frequency (f_0). Nonconstant wavelets with Q increasing from $f_0/\sigma_f=8.5$ to 13 for frequencies from 30 to 100 Hz (step size 2 Hz) were used. Before averaging, all frequency transformations were performed at the single-trial level. Thus, the resulting total power contains signal components phase locked and non phase locked to the stimulus. The resulting power was baseline corrected for each frequency to obtain the relative signal change: $P(t,f)_{corrected}=100\times(P(t,f) - P_{baseline}(f))/P_{baseline}(f)$. The S2 prestimulus period (− 300 to − 100 ms) served as baseline for all spectral analyses. Grand mean time–frequency representations were computed over all subjects. Additionally, the inter-trial 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 and perfect phase locking to the stimulus, respectively. Similar to ERPs, a repeated measures ANOVA with the factors Congruency (congruent, incongruent) and Region of Interest (9 regions) was used to statistically analyze the total power activity in the gamma-band and the ITC.

Source estimations of gamma-band activity

To reconstruct the neuronal sources of the gamma-band activity, a linear beamforming approach was chosen (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 the other activity that is present in the data. All linear beamforming analyses were performed using the FieldTrip open source toolbox (http://www.ru.nl/fcdonders/fieldtrip). A volume conduction model was derived from the MNI template brain resulting in an anatomically realistic 3-shell model. The leadfield matrix was calculated using the boundary element method (BEM) for each grid point in the brain for a regular 7mm 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 (CSD) matrix. The data was first Fourier transformed using wavelet analysis and subsequently the CSD matrix was calculated between all 126 scalp EEG channels (excluding the EOG channels) for the 40–50 Hz response in each trial, separately for a baseline time point (− 200 ms) and an average over three stimulus time points (120, 150, and 180 ms) for each individual subject. Finally, the power change for each grid point was scaled to decibel according to: $P=10\times(\log_{10}(P_{poststimulus}) - \log_{10}(P_{prestimulus}))$. A paired t-test was performed on the result for each grid point to estimate the signal change in each condition and for the difference between conditions across subjects. Subsequently t-values were transformed to z-scores. All source localizations are given in Talairach coordinates (Talairach and Tournoux, 1988).

Results

Behavioral data

All 22 participants that were included in the analysis showed faster response times for semantically congruent compared to incongruent trials. On a group level, the participants' reaction time analysis revealed a significant priming effect showing shorter reaction times for the congruent condition (mean = 803 ms, SD = 92) as compared to the incongruent condition (mean = 925 ms, SD = 167), $t(21)=6.14$, $p<0.001$. Although no response was required to the task-irrelevant visual stimulus, subjects could, in principle, show a response bias arising from subliminal S1 processing. In 48% of the semantically incongruent trials the auditory stimulus was preceded by a visual stimulus that was, theoretically, response-incompatible with the behavioral response given by the subjects. To rule out the possibility that a response bias could account for the prolongation of reaction times, we compared these for the semantically incongruent condition between trials with response-incompatible (mean = 932 ms, SD = 174) or response-compatible (mean = 927 ms, SD = 164) S1 stimuli. This analysis did not reveal significant differences, $t(21)=0.1$, $p>0.9$. Error rates of the response-compatible (mean = 30.2%, SD = 10.6) and response-incompatible trials (mean = 60.3%, SD = 15.2) within the incongruent condition differed significantly ($t(21)=6.12$, $p<0.001$). Significant differences were observed in the error rates between the congruent (mean = 18.8%, SD = 5.19) and the incongruent condition (mean = 45.7%, SD = 6.06), $t(21)=15.82$; $p<0.001$. An analysis according to signal detection theory (Green and Swets, 1966) revealed significant differences of d' between the congruent (mean = 0.83, SD = 0.29) and the incongruent condition (mean = 0.24, SD = 0.24), $t(21)=7.28$; $p<0.001$.

Effects of semantic congruency on event-related potentials

The ERPs of the crossmodal condition for all regions are shown in Fig. 2A, revealing a centrally distributed auditory N1 at 100 ms and a P2 at 200 ms following auditory stimulus onset. Significant differences between the congruent and incongruent condition as indicated by running t-tests started at 250 ms following auditory stimulus onset. The difference wave (incongruent–congruent condition) of the ERPs showed a negative peak around 300 ms. The ANOVA in the

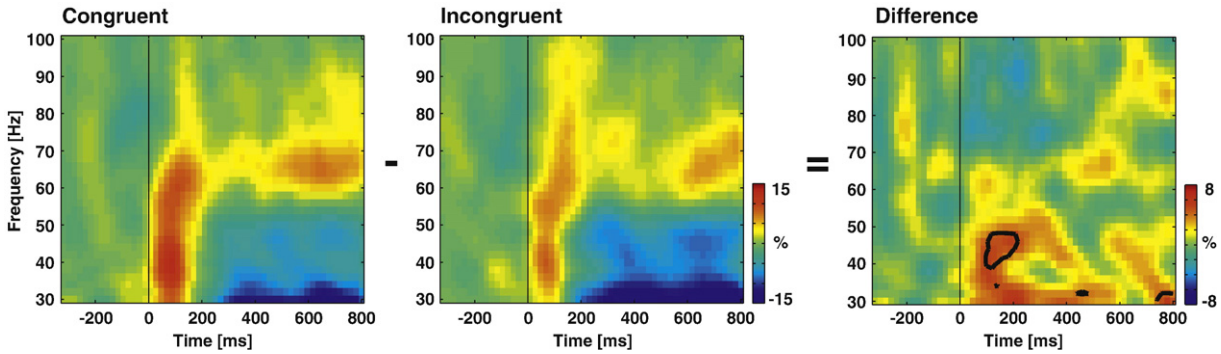

Fig. 4. Time–frequency plots of total gamma-band activity in response to auditory targets (S2) at the medial central ROI. The plots show total oscillatory activity (expressed as percent change relative to baseline) following semantically congruent (left) or incongruent (middle) primes, as well as the difference between the two conditions (right). The contour line in the plot to the right indicates significant differences between the conditions ($p<0.01$).

250–350 ms time window revealed significant main effects of Congruency, $F(1, 21)=67.99$, $p<0.001$, $e=.76$, and Region of Interest, $F(1, 21)=13.4$, $p<0.001$, $e=0.39$, and a significant Congruency × Region of Interest interaction, $F(1, 21)=8.31$, $p<0.001$, suggesting a difference in stimulus processing due to semantic matching. The topographies (Fig. 2B) of the ERPs in the congruent and the incongruent condition revealed bilaterally distributed occipito-temporal generators. The topographic distribution of the difference wave showed a central maximum. In order to test differences in the topographic pattern between the two conditions the global dissimilarity index was computed (Lehmann and Skrandies, 1980), where values of 0 indicate identical maps and values of 2 indicate reversed maps. Dissimilarity values in the 250–350 ms window ranged between 0.029 and 0.088. A bootstrapping approach revealed that the scalp topography did not significantly differ between conditions. The corresponding p-values range between 0.93 and 0.999, suggesting that the topographic distribution was similar in both conditions.

Source estimations of event-related potentials

The LAURA source estimate for the ERPs in the 250–350 ms time window (Fig. 3) of both conditions revealed bilateral activity patterns in middle temporal gyrus (MTG, BA 21, peak coordinates incongruent condition: $x=-59$, $y=-29$, $z=0$) and lateral occipital gyrus (BA 19, peak coordinates congruent condition: $x=47$, $y=-58$, $z=-6$), areas which have been reported to be functionally involved in semantic processing of objects. The strongest source level differences were observed in left MTG (BA 21, peak coordinates: $x=-65$, $y=-29$, $z=0$) and, in addition, bilaterally in middle frontal gyrus (MFG, BA 9, peak coordinates: $x=-48$, $y=16$, $z=30$).

Effects of semantic congruency on gamma-band activity

Fig. 4 shows the time–frequency representation of the auditory S2 response at medial central sensors (ROI MC). A peak in gamma-band activity is present in both conditions, reflecting an oscillatory neural response to auditory stimulation with an early latency. The time–frequency representations display total power, comprising both phase locked and non phase locked fractions of the oscillatory activity. An increase in gamma-band power (40–50 Hz) in the medial central region of interest was found for congruent trials compared to incongruent trials in the time window 120–180 ms. A repeated measures ANOVA showed a significant main effect of Congruency, $F(1, 21)=10.73$, $p<0.01$, $e=0.34$, due to enhanced activity in the gamma-band in the congruent compared to the incongruent condition. No significant effect of Region, $F(1, 21)=1.43$, $p=0.19$ and no significant interaction ($F<1$) were observed. In the 120–180 ms time window, the auditory gamma-band response showed a distributed topography involving central and bilateral temporal electrodes (Fig. 5). Analysis of the inter-trial coherence (Fig. 6) in the respective time–frequency window (120–180 ms; 40–50 Hz) showed no differences between the congruent (mean=0.10) and the incongruent (mean=0.11) condition, $F(1,21)=1.23$, $p=0.23$, indicating that the differences observed in total activity are mainly due to a congruency effect on non phase locked gamma-band signals.

Source estimations of gamma-band activity

Source reconstruction of the auditory gamma-band response in the 120–180 ms time window (Fig. 7A) revealed that the strongest activity for the congruent condition was localized in the left fusiform gyrus, inferior temporal gyrus (ITG), and MTG (BA 20 and 21; peak coordinates: $x=-44$, $y=-3$, $z=-26$; z-score=4.33). For the incongruent condition (Fig. 7B) the peak of the distributed sources was located in left inferior frontal gyrus (BA 45 and 47; peak coordinates: $x=-57$, $y=19$, $z=4$; z-score=3.04). The comparison of the source activity between the congruent and incongruent condition (Fig. 7C) revealed the strongest

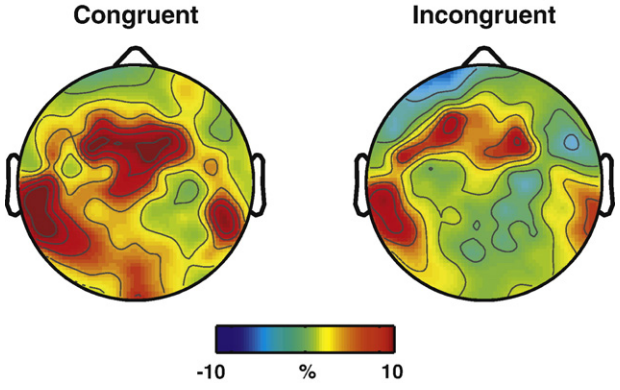

Fig. 5. Topographic maps of the total gamma-band activity (40–50 Hz) in the congruent and incongruent conditions at 120–180 ms.

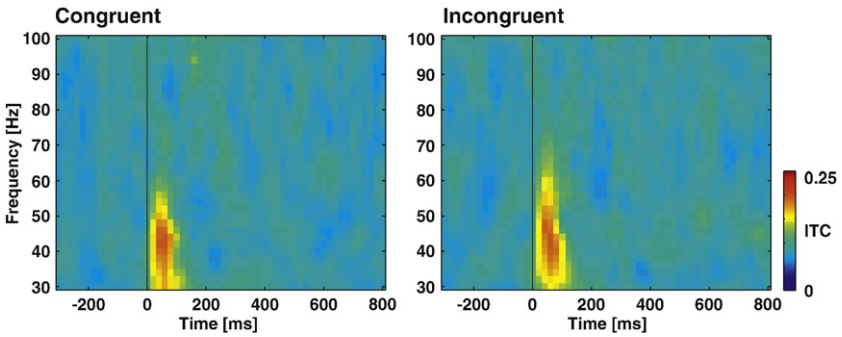

Fig. 6. Inter-trial coherence of the gamma-band activity at the medial central ROI in response to auditory targets (S2) following semantically congruent (left) and incongruent (right) primes.

modulation in left MTG and left ITG (BA 21 and 20; peak coordinates: $x = -44$, $y = -3$, $z = -20$; z-score = 3.57).

Discussion

In the present study we examined crossmodal semantic priming by modulating the semantic congruency between visual–auditory stimulus pairs. A behavioral priming effect was observed, showing facilitated object identification following semantically congruent visual priming, replicating our previous finding (Schneider et al., 2008). Here we find enhanced gamma-band responses following semantically congruent as compared to incongruent stimulus pairs in an early time window (120–180 ms) after onset of the auditory stimulus. Source reconstruction for this gamma-band effect revealed sources in MTG and ITG (BA 21 and 20), regions that have been implicated in crossmodal processing before (Amedi et al., 2005; Beauchamp, 2005). In addition, we find late (250–350 ms) ERP differences between congruent and incongruent stimulus pairs that localize to MTG and MFG (BA 21 and 9).

Multisensory effects on event-related potentials

In our study, the difference wave of the ERPs shows a negative-going deflection for semantically incongruent compared to congruent items peaking between 250 ms and 350 ms, reflecting a crossmodal semantic priming effect. The topography of this ERP effect shows a central distribution. The latency and the topographic distribution of this ERP effect suggest a relationship with the N400, a component which is thought to reflect contextual integration processes (Kutas and Hillyard, 1980; Kutas and Federmeier, 2000). Items which are expected or are related to previously encountered items elicit less negative-going deflections in the ERP around 400 ms as compared to items which are unexpected or unrelated to the semantic context. Semantically congruent and incongruent visually presented words and sounds of objects were presented in a crossmodal priming experiment (Orgs et al., 2006) similar to our study. Differences in the ERPs in response to auditory stimuli in the congruent compared to the incongruent condition started around 200 ms, lasted throughout the classical N400 period, and were interpreted as an N400 effect. Likewise, in a previous study reporting an N400 effect for unrelated compared to related stimulus pairs (spoken words and environmental sounds) was reported with an onset latency of 200 ms in response to sounds (Van Petten and Rheinfelder, 1995). We, thus, conclude that the ERP effects in our experiment are comparable to findings of these conceptual priming studies, which show that contextual integration effects start between 200 ms and 250 ms. Our results suggest that differences in the ERPs with a latency of 250–350 ms may be a signature of a contextual integration process similar to the effects reflected by the N400.

In contrast to other studies that have investigated multisensory integration, we did not find early latency (<200 ms) crossmodal ERP effects, which might be related to the paradigm used in the present study. By presenting complex visual and auditory stimuli simultaneously Molholm et al. (2004) observed an early multisensory effect in the latency range of the visual N1. This finding, however, was interpreted as an effect of co-occurrence of task-relevant features. A semantic congruency effect in the same experiment was reported as late as 450 ms. Investigating unimodal repetition priming, earliest ERP effects were reported at 160 ms in response to auditory stimuli (Murray et al., 2008). In this experiment ERPs were compared in response to initial versus repeated presentation of physically identical sounds. Early crossmodal repetition priming effects in ERPs (100–200 ms) were also reported for written and spoken words, where targets were either repetitions of the prime words or unrelated words (Holcomb et al., 2005). Summarizing these results and integrating with our findings we conclude that crossmodal semantic priming is frequently reflected in ERP differences at longer latencies than crossmodal repetition priming.

Furthermore, electrophysiological findings suggest that unisensory brain responses can be modulated by multisensory memory states (Murray et al., 2004; von Kriegstein and Giraud, 2006; Molhom et al., 2007; Wheeler et al., 2000). Murray et al. (2004) showed that auditory–visual multisensory experiences alter subsequent processing of unisensory visual stimuli as early as 60–140 ms. Similarly, von Kriegstein and Giraud (2006) report activation of multisensory neural networks, which have been shaped by previous associative auditory–visual learning of voices and faces. The spread of activation across modalities reported in these studies could also influence the processing in visual-to-auditory priming. Visual stimuli, although not explicitly associated with auditory stimuli in the experimental situation, but often associatively combined in the everyday environment, may activate multisensory regions and thereby influence auditory signal processing.

Multisensory effects on gamma-band activity

It has been proposed that oscillatory gamma-band responses in the EEG reflect multisensory integration through enhanced coherence among the underlying neuronal populations.

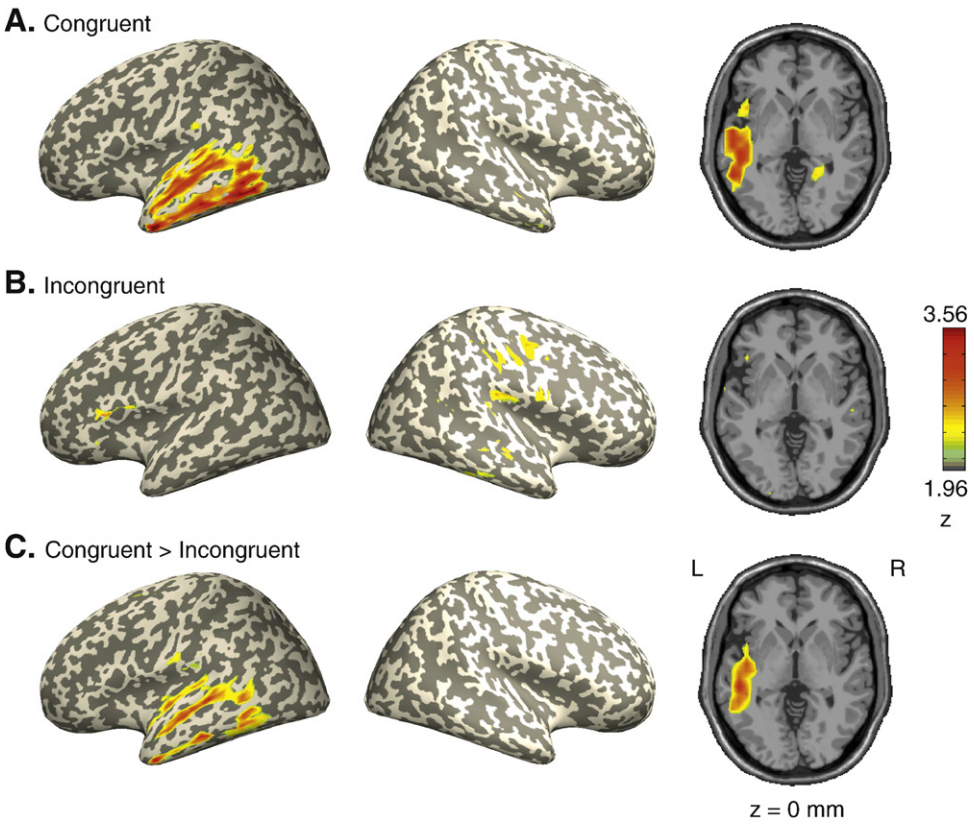

Fig. 7. Source reconstruction of the total gamma-band activity following auditory target onset (S2) in the time window 120–180 ms. The results are plotted on the left and right hemisphere of the MNI brain (left and middle column) and, additionally, on axial slices at $z=0$ (right column). The top two rows (A and B) illustrate the result of a voxelwise dependent sample *t*-test (*t*-values transformed to *z*-scores) between stimulus period and baseline in the congruent and incongruent conditions. The bottom row (C) illustrates the result of the dependent sample *t*-test between the two conditions (congruent > incongruent). The strongest modulations due to semantic congruency are localized in left MTG (BA21).

Enhanced gamma-band activity was reported following matching information in visual and auditory modalities compared with non-matching information (Yuval-Greenberg and Deouell, 2007; Widmann et al., 2007). In a symbol-to-sound-matching paradigm, enhanced gamma-band responses in the EEG were reported for auditory stimuli following matching symbolic visual information (Widmann et al., 2007). In this study simple visual patterns predicted corresponding auditory patterns. Evoked and total power activity in the gamma-band was enhanced following sounds that were congruent to a corresponding visual symbol. Following simultaneous presentation of stimuli representing the same or different objects in the visual and auditory modality enhanced activity in the gamma range was reported peaking at 240 ms poststimulus onset (Yuval-Greenberg and Deouell, 2007). The novelty of the present study lies in the implicit nature of the task , which allows the testing of whether implicit semantic priming processes support a spread of activation across modalities.

Several studies reporting enhanced unisensory auditory gamma-band activity (Debener et al., 2003; Lenz et al., 2007; Widmann et al., 2007) support the match-and-utilization model proposed by Herrmann et al. (2004). According to the match-and-utilization model (MUM) the comparison of bottom-up signals with templates activated in working memory occurs rapidly after stimulus onset and is associated with an early enhanced gamma-band activity if the matching process yields a positive result. Increased activity in the gamma-band was reported for familiar compared to unfamiliar environmental sounds (Lenz et al., 2007), and for auditory target stimuli matching with a template held in working memory (Debener et al., 2003). Although the priming paradigm used in the present study does not involve the production of explicit working memory states, our results, in principle, agree with the match-and-utilization model. The enhanced gamma-band activity observed may reflect a crossmodal semantic matching process that is triggered by the expectation of an upcoming event. The analysis of inter-trial coherence strongly suggests that the reported differences between the two conditions in total gamma-band activity are due to changes in power and not to changes in phase locking to the stimulus. This observation supports the notion that the gamma-band effect reflects a top-down modulation, as it is proposed that changes in phase locking may reflect bottom-up processing, whereas power changes in oscillatory activity may reflect top-down processing (Busch et al., 2006). As predicted by the MUM, ongoing subthreshold oscillations related to the expectation of a stimulus can be augmented by resonance in the respective cell assemblies, if the incoming auditory stimulus matches the expectation. In the case of a semantic mismatch, no such resonance would occur. Our data support this hypothesis and suggest that the enhanced gamma-band activity reflects crossmodal semantic matching processes (Senkowski et al., in press).

Source localization of the ERP and gamma-band effects

As an approach to identifying regions involved in the presumed matching and contextual integration processes we

calculated source estimates for both the early gamma-band and the late ERP effects using linear beamforming and LAURA, respectively. The motivation for choosing two different source analysis techniques for the time and frequency domain is that recent studies have successfully applied linear beamforming for reconstructing the sources of frequency specific activity (Hoogenboom et al., 2006; Bauer et al., 2006; Siegel et al., 2007; Praamstra et al., 2006), whereas LAURA has been shown to be a robust method for source reconstructions of ERPs (Murray et al., 2004, 2006, 2008; Senkowski et al., 2007a).

With respect to the localization of the neuronal populations involved in this crossmodal priming, several scenarios seem possible and are not mutually exclusive (Senkowski et al., in press). First, the condition differences might indicate changes in interactions between distant neuronal assemblies of modality specific areas and reflect informational transfer in a distributed semantic network. Second, the modulations observed might reflect changes in neuronal activity in multisensory convergence sites. Third, the effects may also reflect changes in the interactions between multisensory and unisensory regions.

Source analysis of the ERPs in the latency range between 250 ms and 350 ms revealed sources for both conditions in MTG (BA 21). Furthermore, the maximum for the source level difference between conditions was localized in left MTG (BA 21). This finding is in line with several other studies investigating the sources of contextual integration effects with different methods (for reviewVan Petten and Luka (2006)). Examining the sources of multisensory ERPs in response to auditory–visual objects in motion, Senkowski et al. (2007a) identified a widespread network of occipital, frontal and temporal regions including superior temporal gyrus (STG) and MTG using LAURA. Sources in bilateral occipito-temporal cortex and parahippocampal gyri were reported for the N400 using the variable resolution electromagnetic tomography method (VARETA; Silva-Pereyra et al., 2003). Fell et al. (2004) investigated the N400 by recording intracranially from the medial temporal lobes in epilepsy patients. These results clearly suggest that left temporal regions are of particular importance in contextual integration.

Source analysis of the gamma-band response revealed the strongest sources in left MTG, ITG and fusiform gyrus following semantically congruent stimulation and in left inferior frontal gyrus following semantically incongruent stimulation. The comparison of source activity between the congruent and incongruent condition revealed the strongest modulation in the left MTG and ITG. To our knowledge this is the first study, which reports sources for EEG gamma-band activity in response to auditory inputs using linear beamforming.

Our results are in good agreement with recent neuroimaging studies on visual–auditory integration. Noppeney et al. (2008) report effects of semantic incongruency in a visual-to-auditory priming paradigm in STS and MTG for spoken words and in angular gyrus and intraparietal sulcus for sounds of objects. Response enhancement in STS and MTG was observed when naturalistic auditory and visual objects were presented simultaneously as compared to presentation in a single modality (Beauchamp et al., 2004). In the macaque monkey the superior temporal polysensory area has been reported to be a site of auditory–visual interaction at the level of single cells (Benevento et al., 1977; Bruce et al., 1981). As suggested by Beauchamp et al. (2004) this region could serve as a general-purpose association circuit both within and across the visual and auditory modalities.

Our observation of inferior frontal activation in response to semantically incongruent stimulation might reflect the detection of mismatching information. Inferior frontal gyrus is supposed to play a functional role, for example, in the detection of auditory pattern changes, as enhanced gamma-band activity was observed in this region in a magnetoencephalography study using a visual–auditory McGurk type paradigm (Kaiser et al., 2005).

While these studies suggest that a network involving predominantly left temporal and frontal regions supports visual–auditory integration, other neuroimaging studies indicate that these areas are also activated by unisensory processing of complex auditory stimuli: Investigating the processing of meaningful as compared to meaningless auditory stimuli revealed differences in activation bilaterally in MTG (and STS) dominantly in the left hemisphere (Lewis et al., 2004, 2005; Zatorre et al., 2004; Maeder et al., 2001). Furthermore, recognition of environmental sounds was impaired in patients with lesions in the left hemisphere STG and MTG, while auditory localization and motion perception were preserved (Clarke et al., 2000, 2002). In an electrophysiological study in which sounds of objects were presented in a discrimination task (Murray et al., 2006), source localization of event-related activity revealed a common network of brain regions including the right STG and MTG as well as the left inferior frontal gyrus.

Conclusion

Our study reveals two different effects on semantic congruency across modalities: an early enhancement of gamma-band activity between 120 ms and 180 ms and a late effect in the event-related activity between 250 ms and 350 ms. Gamma-band activity and ERPs provide complementary information and reflect different processes, as can be concluded from the differences in latency, topography and sources of the two effects. The differences in the ERPs between semantically congruent and incongruent conditions are most likely due to different processing in underlying slow wave generators reflecting phase locked activity up to 30 Hz. The differences in the gamma-band response on the other hand are due to differences in high-frequency oscillatory activity at earlier latencies not phase locked to the stimulus. However, source reconstruction of gamma-band activity and ERPs also revealed striking similarities in the networks involved. Both electrophysiological measures show modulations in MTG and different cortical networks are activated depending on whether an auditory stimulus is semantically congruent or incongruent to the visual environment presented. While regions in left MTG and left ITG seem to be activated strongly by matching information, inferior frontal regions may be more involved in the processing of mismatched information. The present results support the notion that neural signals in the gamma-band may be important for multisensory processing (Senkowski et al., in press). The question of whether changes in coherence between distant neuronal assemblies in the gamma-frequency range are involved in multisensory integration remains to be investigated in future research.

Acknowledgments

We thank K. Saha for the acquisition of participants and help with data recording. We also thank J. Hipp, M. Siegel, J. Thorne, D. Senkowski and I. Schepers for their helpful discussions. This research was supported by grants from the European Union

(A.K.E., projects IST 027268, NEST 043457), the Volkswagen Foundation (A.K.E.), the BMBF (A.K.E.), the Forschungsförderungsfonds of the University Medical Center Hamburg-Eppendorf (S.D.), and the G.A.Lienert Foundation (T.R.S.). The Cartool software (http://brainmapping.unige.ch/Cartool.htm) has been programmed by Denis Brunet, from the Functional Brain Mapping Laboratory, Geneva, Switzerland, and is supported by the Center for Biomedical Imaging (CIBM) of Geneva and Lausanne.

References

Amedi, A., von Kriegstein, K., van Atteveldt, N.M., Beauchamp, M.S., Naumer, M.J., 2005. Functional imaging of human crossmodal identification and object recognition. Exp. Brain Res. 166, 559–571.

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.

Beauchamp, M.S., 2005. See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. Curr. Opin. Neurobiol., 15, 145–153.

Beauchamp, M.S., Lee, K.E., Argall, B.D., Martin, A., 2004. Integration of auditory and visual information about objects in superior temporal sulcus. Neuron 41, 809–823.

Benevento, L.A., Fallon, J., Davis, B.J., Rezak, M., 1977. Auditory–visual interaction in single cells in the cortex of the superior temporal sulcus and the orbital frontal cortex of the macaque monkey. Exp. Neurol. 57, 849–872.

Bruce, C., Desimone, R., Gross, C.G., 1981. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. J. Neurophysiol. 46, 369–384.

Buckner, R.L., Goodman, J., Burock, M., Rotte, M., Koutstaal, W., Schacter, D., 1998. Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. Neuron 20, 285–296.

Busch, N.A., Debener, S., Kranczioch, C., Engel, A.K., Herrmann, C.S., 2004. Size matters: effects of stimulus size, duration and eccentricity on the visual gamma-band response. Clin. Neurophysiol. 115, 1810–1820.

Busch, N.A., Schadow, J., Fründ, I., Herrmann, C.S., 2006. Time–frequency analysis of target detection reveals an early interface between bottom-up and top-down processes in the gamma-band. NeuroImage 29, 1106–1116.

Clarke, S., Bellmann, A., Meuli, R.A., Assal, G., Steck, A.J., 2000. Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. Neuropsychologia 38, 797–807.

Clarke, S., Bellmann, T.A., Maeder, P., Adriani, M., Vernet, O., Regli, L., 2002. What and where in human audition: selective deficits following focal hemispheric lesions. Exp. Brain Res. 147, 8–15.

Debener, S., Herrmann, C.S., Kranczioch, C., Gembris, D., Engel, A.K., 2003. Top-down attentional processing enhances auditory evoked gamma band activity. NeuroReport 14, 683–686.

Debener, S., Makeig, S., Delorme, A., Engel, A.K., 2005. What is novel in the novelty oddball paradigm? Functional significance of the novelty P3 event-related potential as revealed by independent component analysis. Brain Res. Cogn. Brain Res. 22, 309–321.

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.

Dobbins, I.G., Schnyer, D.M., Verfaellie, M., Schacter, D.L., 2004. Cortical activity reductions during repetition priming can result from rapid response learning. Nature 428, 316–319.

Engel, A.K., Konig, P., Singer, W., 1991a. Direct physiological evidence for scene segmentation by temporal coding. Proc. Natl. Acad. Sci. U. S. A. 88, 9136–9140.

Engel, A.K., Kreiter, A.K., Konig, P., Singer, W., 1991b. Synchronization of oscillatory neuronal responses between striate and extrastriate visual cortical areas of the cat. Proc. Natl. Acad. Sci. U. S. A. 88, 6048–6052.

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

Fell, J., Dietl, T., Grunwald, T., Kurthen, M., Klaver, P., Trautner, P., Schaller, C., Elger, C.E., Fernandez, G., 2004. Neural bases of cognitive ERPs: more than phase reset. J. Cogn. Neurosci. 16, 1595–1604.

Grave de Peralta Menendez, R., Gonzalez Andino, S., Lantz, G., Michel, C.M., Landis, T., 2001. Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. Brain Topogr. 14, 131–137.

Grave de Peralta Menendez, R., Murray, M.M., Michel, C.M., Martuzzi, R., Gonzalez Andino, S.L., 2004. Electrical neuroimaging based on biophysical constraints. Neuroimage 21, 527–539.

Gray, C.M., Konig, P., Engel, A.K., Singer, W., 1989. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. Nature 338, 334–337.

Green, D.M., Swets, J.A., 1966. Signal Detection Theory and Psychophysics. John Wiley and Sons, New York.

Grill-Spector, K., Henson, R., Martin, A., 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn. Sci. 10, 14–23.

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, M.M., 2002. Effects of picture repetition on induced gamma band responses, evoked potentials, and phase synchrony in the human EEG. Brain Res. Cogn. Brain Res. 13, 377–392.

Gruber, T., Müller, M.M., 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, J.S., 1991. Significance testing of difference potentials. Psychophysiology 28, 240–244.

Herrmann, C.S., Munk, M.H., Engel, A.K., 2004. Cognitive functions of gamma-band activity: memory match and utilization. Trends Cogn. Sci. 8, 347–355.

Hine, J., Debener, S., 2007. Late auditory evoked potentials asymmetry revisited. Clin. Neurophysiol. 118, 1274–1285.

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

Hoogenboom, N., Schoffelen, J.M., Oostenveld, R., Parkes, L.M., Fries, P., 2006. Localizing human visual gamma-band activity in frequency, time and space. NeuroImage 29, 764–773.

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.

Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207, 203–205.

Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends Cogn. Sci. 4, 463–470.

Laurienti, P.J., Kraft, R.A., Maldjian, J.A., Burdette, J.H., Wallace, M.T., 2004. Semantic congruence is a critical factor in multisensory behavioral performance. Exp. Brain Res. 158, 405–414.

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

Lenz, D., Schadow, J., Thaerig, S., Busch, N.A., Herrmann, C.S., 2007. What's that sound? Matches with auditory long-term memory induce gamma activity in human EEG. Int. J. Psychophysiol. 64, 31–38.

Lewis, J.W., Wightman, F.L., Brefczynski, J.A., Phinney, R.E., Binder, J.R., DeYoe, E.A., 2004. Human brain regions involved in recognizing environmental sounds. Cereb. Cortex 14, 1008–1021.

Lewis, J.W., Brefczynski, J.A., Phinney, R.E., Janik, J.J., DeYoe, E.A., 2005. Distinct cortical pathways for processing tool versus animal sounds. J. Neurosci. 25, 5148–5158.

Maeder, P.P., Meuli, R.A., Adriani, M., Bellmann, A., Fornari, E., Thiran, J.P., 2001. Distinct pathways involved in sound recognition and localization: a human fMRI study. NeuroImage 14, 802–816.

Molholm, S., Ritter, W., Javitt, D.C., Foxe, J.J., 2004. Multisensory visual–auditory object recognition in humans: a high-density electrical mapping study. Cereb. Cortex 14, 452–465.

Molholm, S., Martinez, A., Shpaner, M., Foxe, J.J., 2007. Object-based attention is multisensory: co-activation of an object's representations in ignored sensory modalities. Eur. J. Neurosci. 26, 499–509.

Murray, M.M., Michel, C.M., Grave de Peralta, R., Ortigue, S., Brunet, D., Gonzalez Andino, S., Schnider, A., 2004. Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. NeuroImage 21, 125–135.

Murray, M.M., Camen, C., Gonzalez Andino, S.L., Bovet, P., Clarke, S., 2006. Rapid brain discrimination of sounds of objects. J. Neurosci. 26, 1293–1302.

Murray, M.M., Camen, C., Spierer, L., Clarke, S., 2008. Plasticity in representations of environmental sounds revealed by electrical neuroimaging. NeuroImage 39, 847–856.

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

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

Praamstra, P., Kourtis, D., Kwok, H.F., Oostenveld, R., 2006. Neurophysiology of implicit timing in serial choice reaction-time performance. J. Neurosci. 26, 5448–5455.

Schadow, J., Lenz, D., Thaerig, S., Busch, N.A., Fründ, I., Herrmann, C.S., 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, T., Engel, A.K., Debener, S., 2008. Multisensory identification of natural objects in a two-way crossmodal priming paradigm. Exp. Psychol. 55, 121–131.

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., Saint-Amour, D., Kelly, S.P., Foxe, J.J., 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, C.S., Woldorff, M.G., 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, T.R., Foxe, J.J., Engel, A.K., in press. Crossmodal binding through neural coherence: implications for multisensory processing. Trends in Neurosciences.

Siegel, M., Donner, T.H., Oostenveld, R., Fries, P., Engel, A.K., 2007. High-frequency activity in human visual cortex is modulated by visual motion strength. Cereb. Cortex 17, 732–741.

Silva-Pereyra, J., Rivera-Gaxiola, M., Aubert, E., Bosch, J., Galan, L., Salazar, A., 2003. N400 during lexical decision tasks: a current source localization study. Clin. Neurophysiol. 114, 2469–2486.

Singer, W., Gray, C.M., 1995. Visual feature integration and the temporal correlation hypothesis. Annu. Rev., Neurosci. 18, 555–586.
Talairach, J., Tournoux, P., 1988. Co-planar Stereotaxic Atlas of the Human Brain. Thieme Medical Publishers, New York.
Tallon-Baudry, C., Bertrand, O., 1999. Oscillatory gamma activity in humans and its role in object representation. Trends Cogn. Sci. 3, 151–162.
Van Petten, C., Rheinfelder, H., 1995. Conceptual relationships between spoken words and environmental sounds: event-related brain potential measures. Neuropsychologia 33, 485–508.
Van Petten, C., Luka, B.J., 2006. Neural localization of semantic context effects in electromagnetic and hemodynamic studies. Brain Lang. 97, 279–293.
Van Veen, B.D., 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.
Varela, F., Lachaux, J.P., Rodriguez, E., Martinerie, J., 2001. The brainweb: phase synchronization and large-scale integration. Nat. Rev., Neurosci. 2, 229–239.
von der Malsburg, C., Schneider, W., 1986. A neural cocktail-party processor. Biol. Cybern. 54, 29–40.
von Kriegstein, K., Giraud, A.L., 2006. Implicit multisensory associations influence voice recognition. PLoS Biol. 4, e326.
Wheeler, M.E., Petersen, S.E., Buckner, R.L., 2000. Memory's echo: vivid remembering reactivates sensory-specific cortex. Proc. Natl. Acad. Sci. U. S. A. 97, 11125–11129.
Widmann, A., Gruber, T., Kujala, T., Tervaniemi, M., Schroger, E., 2007. Binding symbols and sounds: evidence from event-related oscillatory gamma-band activity. Cereb. Cortex. 17, 2696–2702.
Wiggs, C.L., Martin, A., 1998. Properties and mechanisms of perceptual priming. Curr. Opin. Neurobiol. 8, 227–233.
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.
Zatorre, R.J., Bouffard, M., Belin, P., 2004. Sensitivity to auditory object features in human temporal neocortex. J. Neurosci. 24, 3637–3642.