Multisensory conflict modulates the spread of visual attention across a multisensory object

Ulrike Zimmer [a,*], Kenneth C. Roberts [a], Todd B. Harshbarger [b], Marty G. Woldorff [a,c]

[a] Center for Cognitive Neuroscience, Duke University, USA
[b] Brain Imaging and Analysis Center, Duke University Medical Center, USA
[c] Department of Psychiatry, Duke University, USA

ARTICLE INFO

ABSTRACT

Spatial attention to a visual stimulus that occurs synchronously with a task-irrelevant sound from a different location can lead to increased activity not only in the visual cortex, but also the auditory cortex, apparently reflecting the object-related spreading of attention across both space and modality (Busse et al., 2005). The processing of stimulus conflict, including multisensory stimulus conflict, is known to activate the anterior cingulate cortex (ACC), but the interactive influence on the sensory cortices remains relatively unexamined. Here we used fMRI to examine whether the multisensory spread of visual attention across the sensory cortices previously observed will be modulated by whether there is conceptual or object-related conflict between the relevant visual and irrelevant auditory inputs. Subjects visually attended to one of two lateralized visual letter-streams while synchronously occurring, task-irrelevant, letter sounds were presented centrally, which could be either congruent or incongruent with the visual letters. We observed significant enhancements for incongruent versus congruent letter–sound combinations in the ACC and in the contralateral visual cortex when the visual component was attended, presumably reflecting the conflict detection and the need for boosted attention to the visual stimulus during incongruent trials. In the auditory cortices, activity increased bilaterally if the spatially discordant auditory stimulation was incongruent, but only in the left, language-dominant side when congruent. We conclude that a conflicting incongruent sound, even when task-irrelevant, distracts more strongly than a congruent one, leading to greater capture of attention. This greater capture of attention in turn results in increased activity in the auditory cortex.

© 2010 Elsevier Inc. All rights reserved.

Introduction

Attentional selectivity, the ability to select one particular source of information to guide action while ignoring others that are irrelevant to the current behavioral goal, is a critical cognitive function for successful navigation in the world. Our attentional selection is enhanced when the behaviorally relevant stimuli belong to the same object. Unisensory visual studies of object-based attentional selection have shown an apparently automatic spread of visual attention through a visual cued object (Egly et al., 1994; Abrams and Law, 2000). Correspondingly, an fMRI study has reported higher activation in human visual cortex representing locations in a cued object relative to equidistant locations on another uncued object (Müller and Kleinschmidt, 2003).

Based on these results, in our recent study, (Busse et al., 2005) we asked if visual attention can spread cross-modally to encompass other parts of a multisensory object. In that study, subjects attended to one

of two lateralized streams consisting of rapidly presented simple visual stimuli. Half of these laterally presented visual stimuli were synchronously accompanied by a centrally presented, task-irrelevant tone. Activity associated with the central auditory stimuli occurring synchronously with an attended versus an unattended lateral visual stimulus was extracted by a set of subtraction contrasts. The EEG analysis revealed a late, sustained, frontally distributed pattern of brain activity for the central tones that occurred in the context of an attended relative to an unattended visual stimulus, an activation that resembled attention-related enhancements seen at earlier latencies during intra-auditory selective attention. The corresponding fMRI contrasts showed that this object-related attention effect included enhanced activity in auditory cortex. The pattern of results strongly suggested that attention to the visual modality can spread to encompass simultaneously occurring signals from the auditory modality, even when the auditory stimuli are task-irrelevant and from a different location.

In the Busse et al. (2005) study, the multisensory spread of visual attention across the sensory cortices during attended visual stimulation was revealed with visual–auditory stimulus combinations that were essentially neutral — i.e., they were neither related nor

* Corresponding author. Center for Cognitive Neuroscience, LSRC Rm B242, Box 90999, Duke University, NC-27708 Durham, USA. Fax: +1 919 681 0815.
E-mail address: ulrike.zimmer@duke.edu (U. Zimmer).

unrelated. In the present work, we asked whether this 'spread-of-attention' activity would be modulated by object-related conflict between the relevant visual input and the irrelevant auditory one. To produce object-related multisensory conflict, the stimulus inputs consisted of task-relevant, laterally presented, visual letters ("A" or "X") to be discriminated that could be accompanied by simultaneous, task-irrelevant spoken letters presented centrally (also "A" or "X"). If the visual letter was an "A", the spoken letter could be either congruent ("A") or incongruent ("X"). In previous studies of such multisensory conflict, all the stimuli were typically presented exclusively from the central position, and these have indicated that reading of written letters can be impeded by incongruent letter–sounds and facilitated by congruent ones (e.g. van Atteveldt et al., 2007). Here, the question was whether and by how such multimodal object-related conflict might modulate the cross-modal 'spread-of-attention' activity.

When attention spreads across a multisensory object whose sensory components conflict in some way, we might expect to see increased activity in both the ACC and the visual cortex for incongruent compared to congruent letter/sound combinations. Such effects would be expected for several reasons. First, previous neuroimaging studies have implicated the ACC as being involved in object-related conflict processing (Van Veen and Carter, 2005; Weissman et al., 2003, among others). Secondly, in a previous multisensory cueing fMRI study in which the relevant visual and irrelevant auditory letter stimuli were simultaneously presented from a central location (Weissman et al., 2004), the results indicated that conflict induced by co-occurring incongruent auditory letters led to increased activity in both the ACC and the visual cortex (Weissman et al., 2004). The authors concluded that during conflict between the modalities, attention toward the relevant input is enhanced by boosting activity in the sensory cortex of the task-relevant modality. In our present study, given that the visual letters were relevant and auditory letters to be ignored, we predicted activity in the visual cortex would be enhanced during conflicting stimulation to help counter the auditory semantic distraction. In the current case, we could assess whether such incongruency-related activity patterns in the ACC and visual cortex will occur even during the multisensory spreading of attention across space.

An additional important question investigated here, however, is whether and how the 'spread-of-attention' activity in the auditory cortex would differ for congruent versus incongruent semantic objects. On the one hand, there is evidence (e.g. Molholm et al., 2007) that visual attention would only spread cross-modally if both modalities involved were associated with the same object, and not if they were linked to different objects. In a recent EEG study, Molholm et al. (2007) presented a central stream of alternating pictures and sounds to subjects, with a task to detect a specific picture while ignoring the sounds. Results indicated an increased auditory ERP negativity beginning around 200 ms when the response to the ignored sound corresponded to the attended visual object (e.g. barking sound and dog picture) relative to when the sound that was not a feature of the attended visual object (e.g. barking sound and guitar picture). Thus, the object-based spread of visual attention to the unattended auditory modality was greater when the visual and auditory semantic features matched (Molholm et al., 2007; also see Fiebelkorn et al., 2010).

On the other hand, a task-irrelevant sound that is to be ignored might be more distracting when incongruent with a relevant visual stimulus. Distracting stimuli are known to capture attention (e.g. Sabri et al., 2006; Watkins et al., 2007; cf. Dalton and Spence, 2007 for multisensory distraction). For example, a combined EEG/fMRI auditory study revealed that the more an irrelevant auditory deviant was perceived as distracting, the higher the elicited activity in the auditory cortex (Sabri et al., 2006; see also Watkins et al., 2007).Thus, for the auditory cortex, two opposing predictions could be made. One

prediction hypothesizes that there would be greater 'spreading-of-attention' activity from the attended visual component to the synchronous task-irrelevant auditory component when they are congruent, due explicitly to their congruence engendering increased activity. Enhanced activity for congruent trials relative to incongruent ones could also be because as soon as the incongruency is detected the brain would respond by quickly sending inhibitory signals to suppress its processing. In contrast, if a conflicting incongruent sound (even when task-irrelevant) distracts more strongly than a congruent one, it may trigger a greater capture of attention, thereby leading to more activity in auditory cortex.

In summary, we specifically asked how the 'spread-of-attention' from an attended visual stimulus to a synchronous task-irrelevant auditory stimulus arising from another location would be modulated differentially as a function of whether these components conflicted. Accordingly, we presented congruent and incongruent letter–sound combinations while subjects visually attended one of two lateralized visual letter-streams and ignored the synchronously occurring task-irrelevant central sounds, which could be either semantically congruent or incongruent. Analyses of the interactions of attention by congruency would then provide insight into the mechanisms of the multisensory 'spread-of-attention' process, and into multisensory feature processing interactions more generally.

Materials and methods

Subjects

Sixteen healthy, right-handed subjects (ages 18–35 years; equal numbers of males and females) participated in the experiment. After receiving an explanation of the procedures, all subjects gave written informed consent according to a protocol approved by the Institutional Review Board of Duke University.

Paradigm

Sequences of lateralized visual letters ("A" or "X") were presented to the left and right visual fields (see Fig. 1) while subjects covertly attended to a designated side in each run to discriminate the letter on each trial. On two-thirds of the trials, the lateralized visual letter was accompanied by a synchronous, centrally presented, task-irrelevant, spoken letter (also "A" or "X", male voice) that could be either congruent or incongruent with the lateralized visual one. For each event, both the side of the visual letter (left/right on the screen) and whether it was or was not combined with a task-irrelevant, centrally presented letter–sound were randomized and unpredictable. The side to which the subject was instructed to attend was blocked by run, and every subject completed five 'attend-left' runs and five 'attend-right' runs.

To be able to extract brain activity due to the auditory part of the multisensory stimuli, we included purely visual letter stimuli ("A" and "X", with no accompanying auditory component), the responses to which could be subtracted from the corresponding multisensory response, analogous to the approach used in Busse et al. (2005) with simple unrelated stimuli. The duration of both the auditory and the visual stimuli was 250 ms. To ensure that attention remained tightly focused on the lateralized visual stream, visual stimulus events were presented at the relatively rapid rate of 625 ms, with an MR acquisition rate (i.e., TR) set at 1.25 s. To accomplish this, the congruent, incongruent, and pure visual "A" and "X" stimuli were presented time-locked to the beginning of the TRs (in random order and with equal likelihood), with task-irrelevant visual letters ('H's') being presented as visual 'filling' stimuli on the half-TR. In addition, 10% of the time points on both the full-TR and the half-TR were 'null-events' (i.e., events without any stimulus), which were used to provide some jitter of the stimulus-onsets. For the main on-TR stimuli
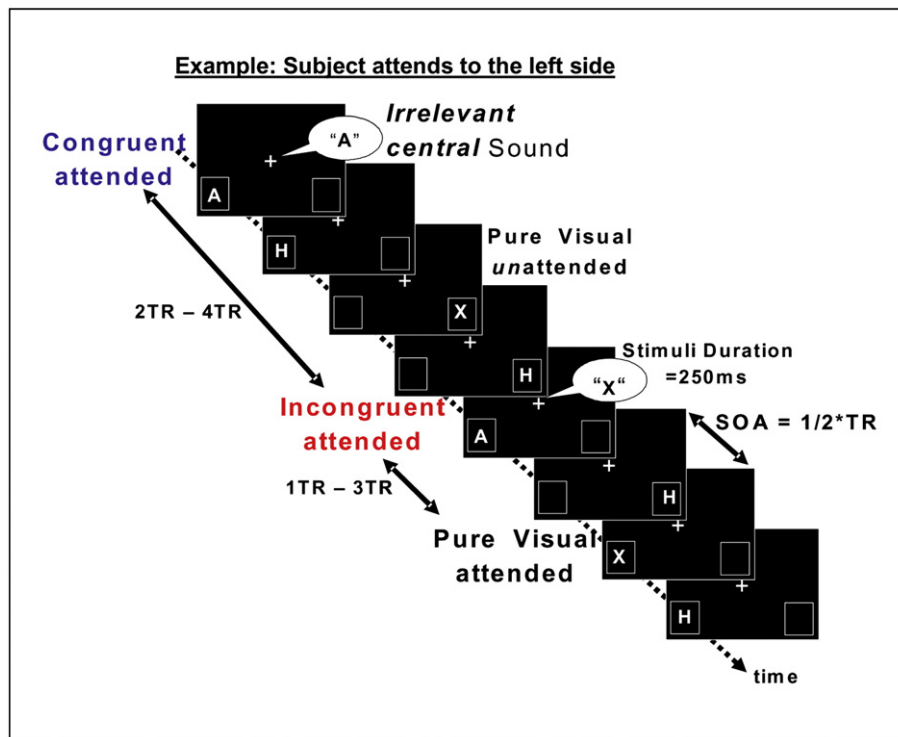
Fig. 1. Task paradigm. An example of stimulus sequence is shown for runs in which subjects attended to the left side. Two visual letter-streams were presented randomly to the left and right visual fields, including two task-relevant letters ("A", "X", each 30% of the total events). The task of the subject was to fixate on the central white cross and to detect "A" versus "X" on the designated (i.e. attended) side only. Two-thirds of the lateral visual letters (both attended and unattended) were accompanied by a simultaneous, centrally presented, spoken letter sound, 50% of which were incongruent with the visual letter and 50% congruent. Subjects were instructed to ignore the central spoken letters.

of interest, the stimulus pseudo-randomization was constrained so as to prevent two multisensory stimuli from occurring in consecutive TRs. This was done to minimize the possibility of saturation of the functional signal in auditory cortex, and to avoid any adaptation to the conditions of either congruency or incongruency (cf. Egner and Hirsch, 2005 for the visual cortex).

Each letter was presented 9.5° to the left or right of the fixation cross, and vertically 4° below fixation. This lateralized lower-visual-field positioning would be expected to evoke low-level brain activity in contralateral visual cortex dorsal to the calcarine sulcus. The letters were presented in two rectangular boxes (size 3.6°×4.8°). These two laterally located boxes remained on continuously over the entire run, serving as attentional anchors to assist the subject in maintaining a strong attentional focus on the spatial position of the attended visual letter stream.

At the beginning of each run, subjects fixated on a central fixation cross and directed their attention to the anchor box on the designated side of their visual field. The subjects' task was to attend covertly to the stream of visual letters presented in the box on the designated side and to press one of two buttons with their right index finger when they saw the letter "A" appear there and the other button when they saw an "X". Subjects were instructed to ignore the visual stimuli on the other side of their visual field and to also ignore any sounds.

Image acquisition

Imaging was carried out in a 3 T whole-body scanner (GE Signa EXCITE HD system). Structural images for each participant were collected using a T1-weighted gradient-echo sequence with FOV 256 mm×256 mm×176 mm and a resolution of 1 mm×1 mm× 1 mm. Functional BOLD (blood oxygenation level-dependent) contrast was obtained using a spiral imaging sequence with SENSE acceleration. The acquisition consisted of 32 transverse slices; thereby providing coverage of the whole cerebral cortex, acquired with a

repetition time (TR) of 1.25 s. The in-plane resolution was 3 mm× 3 mm, with a slice thickness of 4 mm.

Data analysis

Psychophysical data

Only trials for which the behavioral responses occurred between 200–1000 ms after target presentation were considered for further analysis. Reaction times (RTs) for correct letter discrimination of the attended-stream visual letters were computed separately for the congruent, incongruent, and pure-visual trial types. Analyses of variance (ANOVAs) and subsequent paired *t*-tests were performed on the RTs between these experimental conditions.

MRI data

The MRI data were analyzed using the software package SPM2 (http://www.fil.ion.ucl.ac.uk). The first four image volumes of each run were discarded to allow for stabilization of longitudinal magnetization, leaving 1620 volumes per subject. Preprocessing included rigid-body transformation (realignment) to correct for head movement. The images were normalized to MNI space (Friston et al., 1995, Mazziotta et al., 1995), using the mean of the functional volumes, and then smoothed with a Gaussian filter of 6 mm full-width at half maximum (FWHM) to increase the signal-to-noise ratio and to facilitate group analyses.

For the analysis of the present study, we decided to use a finite impulse response (FIR) model as we presented stimuli in rather rapid succession (task-relevant stimuli could appear as closely as 1250 ms apart). The FIR has a unique ability to model any HRF, without constraints on the shape of the response. Such a feature becomes even more important when events are presented rapidly. For example, some recent studies with fast event-related presentation found a shift of the peak of the BOLD response from 6 s (the peak of the canonical response used in SPM) to as early as 4–5 s in superior colliculus (Wall

et al., 2009), as well as in visual and/or auditory cortices (Burock et al., 1998; Weissman et al., 2009). Accordingly, the FIR model seemed to be a reasonable choice for the present study.

Statistical inference was based on a random effects approach (Friston et al., 1998), which comprised two steps. First, for each subject, the data were modeled separately for attended left and attended right runs, resulting in two separate design matrices per subject. Each design matrix modeled six event types, which was derived from the crossing of the two factors of attention (att/unatt) and stimulus type (pure/con/inc). For each of these six event types, ten TRs (12.5 s) of the BOLD response were modeled by using a finite impulse response (FIR) model, resulting in a total of 60 regressors in the design matrix. Of the ten modeled time points in the FIR, the peak and the two closest time points to the peak of the bold time course were averaged (for each trial type) to use in a second stage of modeling that treats subject variation as a random effect. The second stage of modeling took as input the six contrast images from the attend-right runs and the six from the attend-left runs, resulting in 12 contrast images per subject, which were then used to construct an ANOVA with subject as a repeated measure ROIs for areas of interest that were established by either an orthogonal functional contrast or from locations described in previous literature. These were then tested for the effects of attention and conflict and their interactions.

Definitions of ROIs

Lateralized functional ROIs of the visual cortex were established by contrasts between contra- and ipsilateral visual stimulation independent of stimulus type (as in, e.g., Macaluso et al., 2000; Zimmer and Macaluso, 2007; see Table 1). Functional ROIs of the left and right auditory cortices in the superior temporal lobe were established by subtracting all pure visual activation trials from all multisensory activation trials (as in, e.g., Weissman et al., 2004: see Table 1). For the ACC, we functionally defined a region of interest by using an F-contrast over all conditions (i.e., all effects of interest) in the entire brain. The voxel-threshold for this F-contrast was set to $p < 0.0001$ (uncorrected), allowing the identification of relatively large numbers of activated voxels in the cognitive division of the ACC. In the resulting F-map, the ROI was centered in a 10 mm sphere around areas of functional activation related to conflict processing and response selection from previous studies (at $x = 6$ mm, $y = 6$ mm, $z = 45$ mm; the centroid of activations in Fan et al., 2008; Weissman et al., 2004; Picard and Strick, 2001, see Table 1). Note, that by forming each ROI, we were interested which of the single voxels in the whole brain were specific for our functional contrast. Thus, we have used a whole brain correction based on the corresponding underlying functional contrast (Table 1).

Testing for areas of attentional spread in each ROI

Across the different regions of the brain, we were interested if the multisensory spreading of attention to the auditory component would be modulated by object-related conflict. We therefore examined in each ROI the differential activity between visually attended multisensory stimuli and visually unattended multisensory stimuli by using an exclusive mask that discounted areas that showed an attentional effect in the case of pure visual stimulation alone. This exclusive mask was fairly liberal, in that it excludes any areas showing an attentional effect of pure visual stimulation reaching an uncorrected p-level of 0.05 (cf. Uncapher and Rugg, 2005; Christensen et al., 2006). Note that this masking does not test for any differences between visual attention effects and multisensory attention effects (http://www.fil.ion.ucl.ac.uk/spm/doc/manual/manual.pdf). In other words, our exclusive mask of voxels that show an attentional effect for pure visual stimuli *does not affect the p-values of the target contrast* (all multisensory versus all visual stimulation, see Table 1). Once areas revealed by this masking were identified in a manner that is irrespective of congruency, subsequent t-tests were necessary to test if the congruent and/or incongruent trial types elicited significantly increased activity relative to any possible residual visual attention effects. Then these effects could be compared between conditions of congruency and incongruency spreading-of-attention in an orthogonal manner. Note that our follow-up tests over the entire masked auditory ROI were aimed to test for attentional differences between stimuli of differential congruency, thus staying orthogonal to the masking.

To test for significance of the multisensory spreading of attention, we extracted the hemodynamic time course using a finite impulse response (FIR) model (10 modeled time points) for each condition in the areas of multisensory attention outside the mask of visual attentional effects (using the MarsBaR toolbox for SPM, http://marsbar.sourceforge.net/). We then averaged over the peak maximum and its two closest time points to reveal the 'percentage signal change'. In all our ROIs (ACC, visual and auditory cortices), this results in TRs 3, 4, and 5 (i.e. 3.75–6.25 s; see the time course of activity in Figs. 2B, 3B and 4B and C) being used to compare the differential responses of the various trial types. These differential responses were tested over the entire ROI as a whole, thus asking if the *entire* ROI (instead of single voxels within the ROI) is significantly affected for a specific comparison. Thus, search-volume corrections for such ROI analyses are not necessary (Poldrack, 2007).

Correlation analyses

To gain further insight into the functional significance of the interaction between conflict and the spreading of attention in the

Table 1
ROI statistics.

ROI	x, y, z (MNI)	Size unmasked (green areas)	Size masked (red areas)	Z-value	p-cor	p-uncor
Anterior cingulate cortex						
ACC (*)	2, 12, 40	144 vox	92 vox	5.92	<0.0001	<0.0001
Visual cortex						
Left ROI	−32, −76,−10	202 vox	68 vox	7.69	<0.0001	<0.0001
	−18, −94,−16	34 vox		6.34	0.014	<0.0001
Right ROI	26, 74, 8	49 vox	NaN	5.32	0.003	<0.0001
Auditory cortex						
Left ROI						
Superior temporal gyrus	−56,−20,4	465 vox	42 vox	5.97	0.003	<0.0001
	−38,−34,14		39 vox	5.69	0.004	<0.0001
Right ROI						
Superior temporal gyrus	64, −24, 6	282 vox	90 vox	6.92	<0.0001	<0.0001

Statistics of ROIs (green and red areas). Note that exclusive masking only excludes voxels showing an attentional effect of pure visual stimulation at an uncorrected p-value of 0.05, but importantly does *not* affect p-values of the functional contrast. (http://www.fil.ion.ucl.ac.uk/spm/doc/manual/manual.pdf). (*) values for the ACC were revealed with an F-test over effects of interest centered on a 10 mm sphere around the anatomical ACC center (x,y,z = 6,6,45, see Fan et al., 2008; Picard and Strick, 2001).
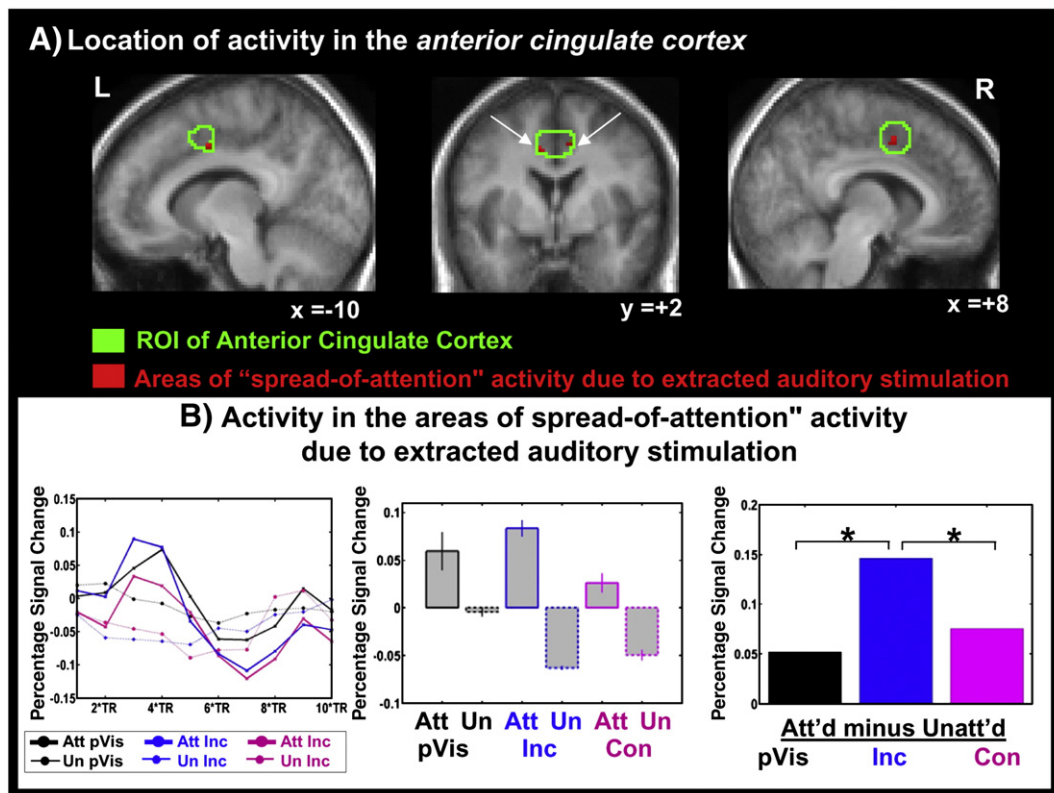
Fig. 2. Effects in the anterior cingulate cortex (ACC). Top panel (A): Green outlines show the conflict-detecting area of the ACC based on the literature (see Materials and methods). Red shading shows the areas of 'multisensory spreading of attention'—areas showing a significant effect of attention in the multisensory trials but not in the pure visual trials. This 'multisensory attentional spreading' was revealed by collapsing over congruent and incongruent multisensory stimuli. Activations are overlaid on the mean of the anatomical images of the 16 subjects. This 'multisensory attentional spreading' region was found to be significant at a corrected threshold of $p = 0.05$. Bottom panel (B): The time course of activity over all areas of multisensory spread (i.e., over all red areas in Fig. 2A; left panel) and the resulting signal bar plots averaged over BOLD peaks, shown separately for the attended-visual and unattended-visual conditions for the different stimulus types (central panel), along with the corresponding attentional difference effect (i.e., 'spreading-of-attention' effect) for the different stimulus types (right panel). Attentional differences were revealed by subtracting the visually unattended signal activity from the visually attended signal activity for each type of stimulation. Within-ROI analyses that were found to be significant at $p = 0.05$ are indicated on the bar graph.

auditory cortex, we correlated this interaction activation across subjects with the behavioral difference in reaction times between congruent and incongruent stimulation. In addition, we tested for a correlation of this interaction activity in auditory cortex with the interaction-related activity in both the ACC and the visual cortex. We used the extracted maximum 'percentage signal change' of TR = 3–5 (see above and see the time course of activity in Figs. 2B, 3B and 4B and C) for a linear regression across subjects to determine the degree of correlation of those conflict-attention BOLD interaction peaks in auditory cortex with the behavioral measures and with the activation in the ACC and visual cortex.

Results

### Behavioral data during fMRI scanning

Subjects were instructed to visually attend to the letters presented on the designated side and to discriminate between the letters "A" and "X" with a button press. Reaction times for correct responses, pooled over the responses to visual letters "A" and "X", were as follows: pure visual presentation 573 ms (SD 61 ms); congruent presentation 541 ms (SD 59 ms), and incongruent presentation 549 ms (SD 59 ms). A repeated-measures ANOVA revealed a highly significant effect of stimulus presentation ($F(2,29) = 29.45$; $p<0.001$). Subsequent paired t-tests indicated that reaction times to pure visual presentation were significantly slower than both congruent ($t_{(15)} = 5.74$; $p<0.001$) and incongruent trials ($t_{(15)} = 5.32$; $p<0.001$), suggesting a multisensory processing facilitation effect of the sound,

possibly due to a general arousal effect (cf. Noesselt et al., 2008). In addition, however, subjects' reaction times differed significantly with the congruency relationship of the sound stimuli, being slower for incongruent versus congruent trials ($t_{(15)} = 3.73$; $p = 0.002$). Thus, while attending the visual letters, subjects appeared to be distracted by an incongruent sound, despite it being completely task-irrelevant and arising from a different location.

fMRI data

Multisensory spread of attention and its interaction with conflict in the ACC

Within the ACC ROI, multisensory spreading of attention due to the auditory stimulus component was found bilaterally in two small regions (Fig. 2A, red areas, see in Table 1 masked size, cf. Materials and methods). The time course of this area of multisensory spreading of attention due to the extracted auditory component in the ACC shows the peak maximum and its two closest time points at TRs 3, 4, and 5 (i.e. 3.75–6.25 s, Fig. 2B left panel). We averaged over these time points to extract the 'percentage signal change' values shown in the signal br plots for this area (Fig. 2B, central panel), which show that activity increased with attention for all the stimulus types (i.e., pure visual, congruent, and incongruent), indicating a general effect of attentional enhancement. The right side of Fig. 2B shows the difference of the attended minus the unattended trials for each stimulus type revealed by subtraction of the activity shown on the left side of Fig. 2B. The difference plots indicate that the 'spreading-of-attention' activity is significantly different for incongruent versus pure
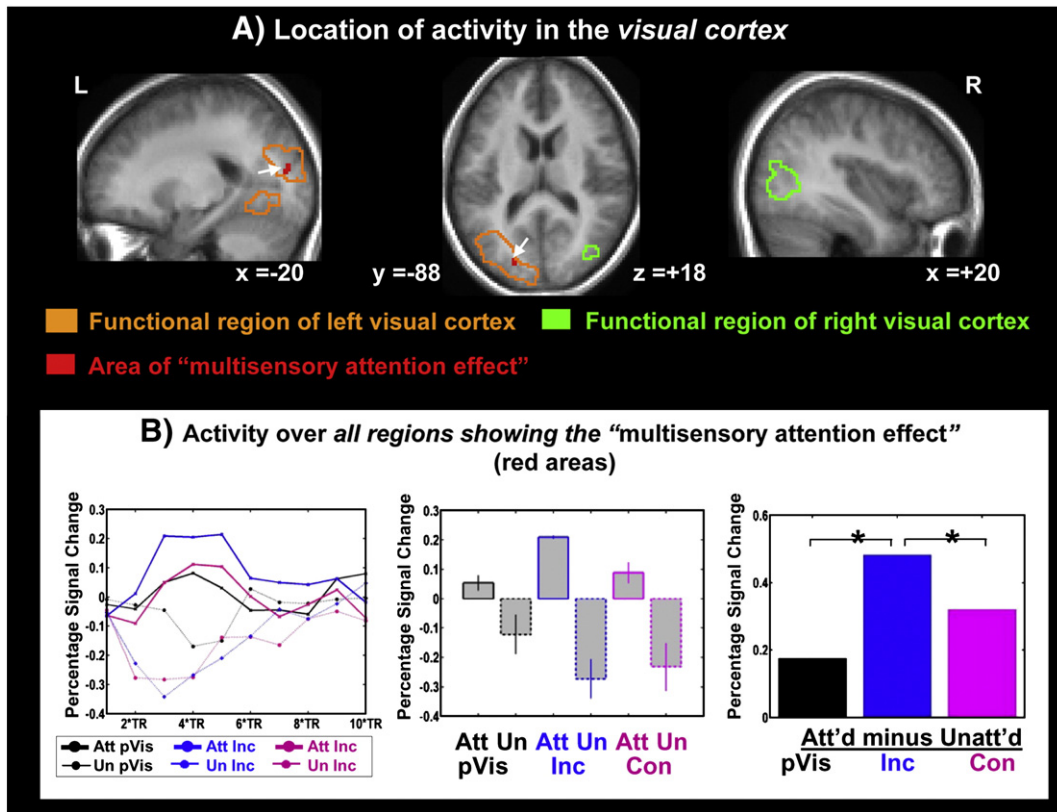
Fig. 3. Effects in visual cortex. Top panel (A): Areas of 'conflict-induced attentional boosting' collapsed over congruent and incongruent trials in the functionally defined regions of visual cortex (red areas in green/orange outlines). Activations are overlaid on the mean anatomical image of all 16 subjects. White arrows note the location of the significant multisensory attentional effects. This 'multisensory attention effect' region was found to be significant at a corrected threshold of $p = 0.05$. Bottom panel (B): The time course of activity over all areas of multisensory attentional enhancement in visual cortex (red area in Fig. 3A; left panel) and the resulting bar plots of activity changes in this area, shown separately for the attended-visual and unattended-visual conditions for the different stimulus types (central panel), along with the corresponding attentional difference effect. Attention-related differences and thresholds were calculated as in Fig. 2.

visual stimulation, as well as different for incongruent versus congruent stimulation. However, the comparison between congruent multisensory trials and pure visual trials did not reach significance (left Fig. 2B and Table 1A). Thus, in the ACC there was only significant spreading of attention for the incongruent, but not congruent condition. This result is consistent with previous literature (Van Veen and Carter, 2005; Weissman et al., 2003 among others), indicating that the ACC shows greater activation for incongruent than for congruent stimulation.

Multisensory attentional spread and interactions in the visual cortex

The functional ROI of the visual cortex (Fig. 3A, green and orange outlines, see unmasked size in Table 1; cf. Materials and methods) was revealed by subtracting all ipsilateral activity from all contralateral activity regardless if evoked by uni- or multisensory stimulation. In these functionally defined areas of the visual cortex, we found multisensory spreading of attention due to the extracted auditory component in the left dorsal visual cortex, corresponding to the position of the letter-boxes below the horizontal meridian in the lower-visual field (red area, Fig. 3A; Table 1 masked area Materials and methods). The time courses of this area for the different conditions are shown in Fig. 3B, left panel. Averaged over these three time points, the bar plots for this 'multisensory attention effect' area (Fig. 3B, central panel) show that activity increased with attention for each stimulus type, thus again reflecting a main effect of attentional enhancement. This effect seems likely to be an attentional boosting of the relevant visual modality resulting from the co-occurence of an auditory stimulus. The difference plots (Fig. 3B right) indicate that the attentional differences for incongruent letter/

sound pairs are significantly increased compared to the attentional differences of pure visual stimulation, presumably reflecting a boosting of activity due to increased visual attention in response to the conflicting stimulus components in the incongruent letter/sound pair. The corresponding comparison for the congruent condition revealed a trend ($p = 0.085$) toward a spreading of attention also in the congruent condition (Fig. 3C right side, Table 2). Attention-related differences were also significantly higher in the incongruent than in the congruent condition (Fig. 3C, left side and Table 2). In conclusion, the specifically enhanced attentional gain for conflicting stimuli seems to support the hypothesis that the distraction due to an ignored, simultaneously presented incongruent sound may be countered by boosting visual processing of the relevant visual letter stimulus.

Multisensory attentional spread and interactions in the auditory cortex

The ROI of the auditory cortex (Fig. 4A, green outlines, Table 1: unmasked size) was revealed by subtracting all pure visual activity from all multisensory activity. The resulting auditory ROI included primary auditory areas (see Penhune et al., 1996; Morosan et al., 2001; Rademacher et al., 2001), as well as posterior and anterior secondary auditory areas.

As mentioned above, the centrally presented auditory stimuli were completely task-irrelevant, and any attention-related differences observed for the extracted auditory responses presumably resulted from a 'spreading of attention' from the visual to the auditory stimulus components. Since the content of the auditory stimulus was spoken letters, we reasoned that these language stimuli might be differentially processed in left and right auditory cortex due to the well-
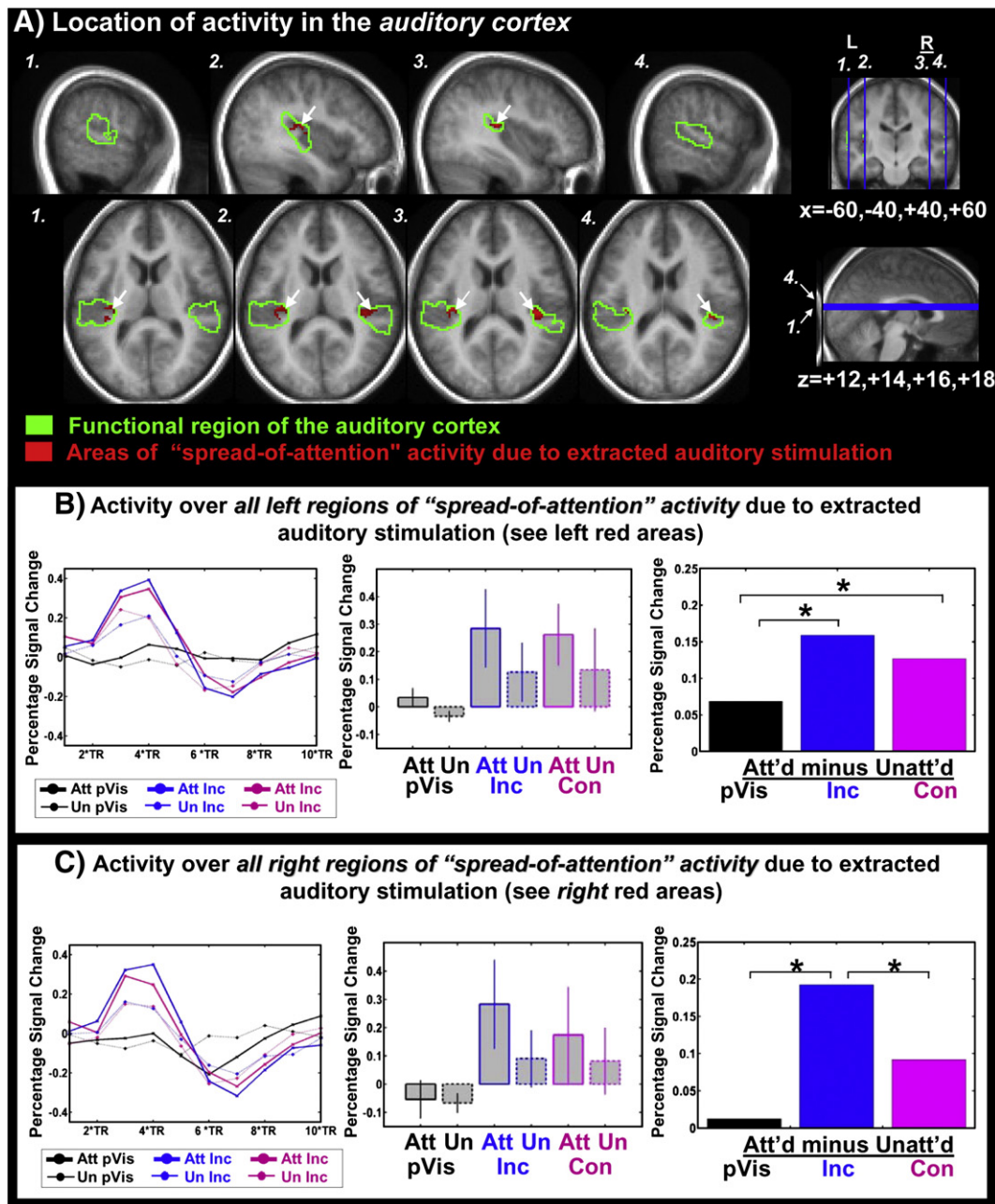
Fig. 4. Effects in auditory cortex. Top panel (A): Areas of 'multisensory spreading of attention' collapsed over congruent and incongruent trials are shown in the functionally defined region of the auditory cortex (red areas in green outlines). Activations are overlaid on the mean anatomical image of the 16 subjects. Within-ROI analyses that were found to be significant at $p = 0.05$ are indicated on the bar graph. Bottom panels (B,C): Multisensory spreading and interaction in (B) left auditory cortex (C) right auditory cortex: The time course of activity over all areas of multisensory spread (red areas in Fig. 3A; left panels in B and C) and the resulting bar plots of activity changes in the areas of interaction, shown separately for the attended-visual and unattended-visual conditions for the different stimulus types (central panels in B and C), along with the corresponding attentional difference effect (i.e., 'spreading-of-attention' effect) for the different stimulus types (right panels in B and C). Attention-related differences and thresholds were calculated as in Fig. 2.

established language dominance of the left auditory cortex (e.g. Dorsaint-Pierre et al., 2006; Tzourio et al., 1998). Therefore, we investigated any effects of multisensory spreading of attention separately for the left and right auditory cortex. The areas of multisensory spreading of attention to the extracted auditory component, revealed by the exclusive masking (see Materials and methods), restricted the ROI's of both left and right auditory cortices to mainly in and near primary auditory cortex, i.e. Heschl's Gyrus (see red areas in Fig. 4A, Table 1: masked size). In left auditory cortex, the time course over the area of multisensory spreading of attention and their attentional differences showed the peak maximum and its two closest time points at TRs 3, 4, and 5 (i.e. 3.75–6.25 s, Fig. 4B left

panel). Averaged over these time points, the signal plots over this area of multisensory spreading of attention and their attentional differences (Fig. 4B) indicated that incongruent and congruent attentional differences were both significantly increased compared to attentional differences for the pure visual stimuli (Fig. 4B). Thus, in the left auditory cortex, activity that reflects a spreading of attention to the task-irrelevant auditory sound was found independently of the congruency of that sound. In contrast, in the right auditory cortex, the activity patterns (Fig. 4C, central panel) indicated that multisensory spreading of attention was significantly greater for the incongruent trials than for the congruent ones, as well as being significantly larger than any residual attention-related differences for

Table 2
Activity changes in areas of multisensory 'spreading of attention'.

| Activity changes in the unified 'spreading-of-attention' regions (red areas in Figs. 2–4)		
	$t_{(15)}$-value	p
Anterior cingulate cortex (ACC)		
Multisensory spread (Con and Inc averaged)	2.213	0.049
Incongruent spread	2.499	0.032
Congruent spread	*0.181*	*0.430*
Difference incongruent vs. congruent spread	2.511	0.026
Visual cortex		
Multisensory spread (Con and Inc averaged)	2.289	0.048
Incongruent spread	2.445	0.036
Congruent spread	*1.446*	*0.085*
Difference incongruent vs. congruent Spread	2.707	0.015
Left auditory cortex		
Multisensory spread (Con and Inc averaged)	2.379	0.040
Incongruent spread	4.242	<0.001
Congruent spread	2.888	0.013
Difference incongruent vs. congruent spread	*0.543*	*0.298*
Right auditory cortex		
Multisensory spread (Con and Inc averaged)	2.445	0.036
Incongruent spread	3.673	0.002
Congruent spread	*0.949*	*0.179*
Difference incongruent vs. congruent spread	3.066	0.008

'Spreading-of-Attention' activity revealed in regions of collapsed multisensory 'spreading-of-attention' and by subsequent within-ROI analyses. Comparisons that do not reach statistical significance are presented in *italics*.

the pure visual trials (Fig. 4C, right panel). Thus, for incongruent trials, 'spreading-of-attention' was observed in both the left and right auditory cortices and was therefore apparently independent of semantic relations in the multisensory components. In contrast, a synchronous congruent letter sound seems only to elicit a spreading of attention when processed in the language-dominant left auditory cortex, suggesting that this multisensory spreading of attention might be facilitated by semantic binding between the multisensory components.

Conflict-related correlations between brain and behavior and between brain areas

The experimental design yielded a behavioral measure of perceptual conflict, namely the reaction time difference between incongruent and congruent multisensory trials, along with multiple brain-activation measures. So far, our results indicated greater spreading-of-attention activity for the incongruent versus congruent letter/sound combination in multiple brain areas (see Figs. 2–4), consistent with the conclusion that an incongruent sound acts as a greater distracter that captures more in-process attentional resources. If this hypothesis is true, then the more the task-irrelevant conflicting sounds capture attention, the greater should be the behavioral effects on reaction time, with a corresponding increase in the difference in 'spreading-of-attention' activity between incongruent and congruent trials. Further, the greater the distracter-related differences in 'spreading-of-attention' activity in the auditory cortices, the greater should be the corresponding effect in the ACC, due to the increased distraction to conflicting input.

These predictions were tested with several correlational analyses across subjects. First, we performed a linear regression of incongruent
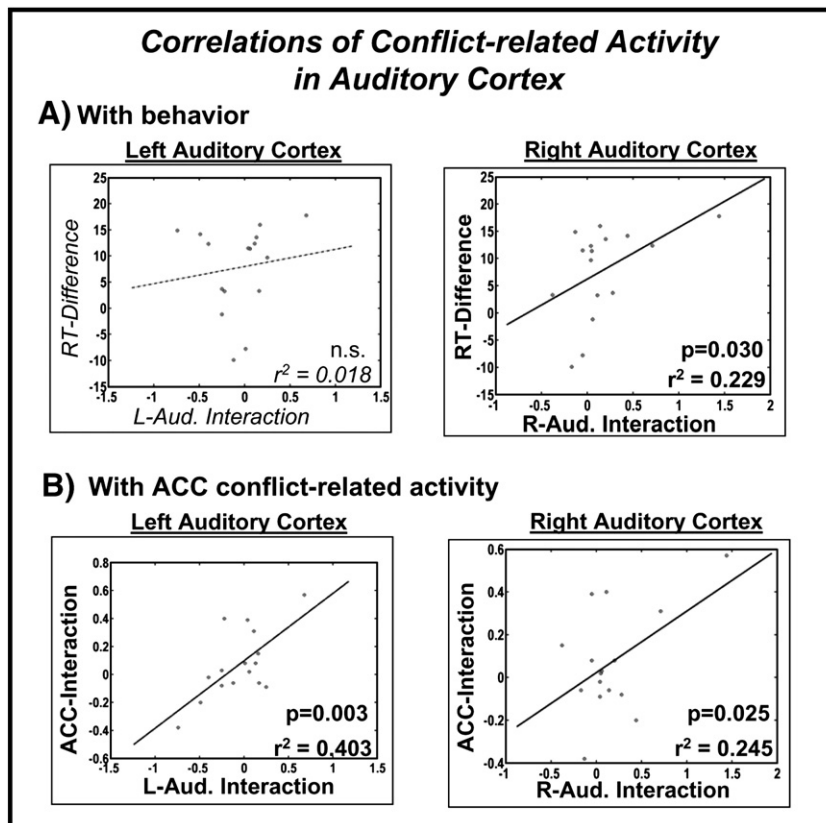
Fig. 5. Correlations of conflict-related-modulation in auditory cortex: with behavioral interference and between brain areas. Panel (A): Correlations with reaction times (RT), shown separately for the left and right auditory cortices. For each subject, the behavioral interference is expressed in reaction time differences to incongruent-minus-congruent letter/sound combinations; the maximum 'percentage signal change' of the BOLD signal is displayed for the interaction of conflict and attention in auditory cortex (single black points). Panel (B): Correlations of observed conflict-related activations in L/R auditory cortex and ACC. For each subject, the maximum 'percentage signal change' of the BOLD signal is displayed for the interaction of conflict and attention in auditory cortex (Left/Right) and that in the ACC (single black points).

versus congruent differences in spreading-of-attention activity in either the right or left auditory cortex and the RT-measurement of stimulus conflict. In the right auditory cortex, this correlation analysis was significant (Fig. 5 left upper panel; $r^2 = 0.23$; $p = 0.03$; $t_{(15)} = 4.2$, one-sided t-test). In the left auditory cortex, this correlation did not reach significance (Fig. 5 lower left panel). Secondly, we tested for correlations of the difference in incongruent versus congruent spreading-of-attention activity between brain regions − namely between the ACC and the right and left auditory cortices. Both of these correlations were found to be significant (Fig. 5 right upper panel; $r^2 = 0.25$; p = 0.025; $t_{(15)} = 4.55$, one-sided t-test; right lower panel; $r^2 = 0.40$; $p = 0.003$; $t_{(15)} = 11.1$, one-sided t-test), suggesting that increased distractor-related activity in auditory cortex for incongruent attended trials is associated with enhanced conflict effects in the ACC.

Discussion

In the present study, we investigated how the 'spread-of-attention' across a multisensory audiovisual object shown in Busse et al. (2005) for simple unrelated stimuli would be modulated when the spatially discordant auditory and visual stimulus components conflicted. In agreement with our predictions, we observed significant enhancements for incongruent versus congruent letter–sound combinations in the ACC and in the contralateral visual cortex when the visual component was attended. In the auditory cortex, our results showed greater spreading of attention for incongruent sounds than for congruent ones. Moreover, the 'spreading-of-attention' activity was found to be significantly present in both auditory cortices if the spatially discordant auditory stimulation was incongruent, but mainly in the left language-dominant auditory cortex if it was congruent. Such a pattern supports a model in which a task-irrelevant sound occurring synchronously with an attended visual event, even if from a different location, tends to distract more strongly than congruent ones, and in turn leads to a greater capture of attention and a greater competition for attentional resources.

Anterior cingulate cortex (ACC)

In the ACC, the 'spread-of-attention' during multimodal object-related conflict resulted in increased activity relative to conditions without conflict, confirming our prediction for this area. In general, results of previous neuroimaging studies have indicated increased activity in the anterior cingulate cortex (ACC) for incongruent multidimensional stimuli compared to the congruent ones (Van Veen and Carter, 2005; Weissman et al., 2003, among others). More specifically, in a recent multisensory cuing paradigm (Weissman et al., 2004) in which subjects had to identify visual letters while ignoring task-irrelevant auditory letter stimulation, both centrally presented at the point of fixation, the anterior cingulate cortex showed enhanced activity when the auditory letter information conflicted with the attended visual source. Furthermore, ACC activity was also shown in a modified Stroop paradigm where the conflicting stimuli were spatially separated (Fan et al., 2003). Here, we found a similar involvement of the ACC, specifically when the visual component of the incongruent multisensory stimuli was attended versus unattended, and thus when visual attention was spreading over both modality and space to the task-irrelevant auditory component.

Visual cortex

Confirming our prediction for the visual cortex, we found increased activity corresponding to a boosting of attention to the relevant modality for incongruent compared to congruent letter/ sound combinations. A recent fMRI study on global versus local conflict (Billington et al., 2008) provided evidence that activity in the visual cortex is enhanced for conflicting versus non-conflicting visual stimulation, and therefore supports processing of the relevant events. In our previous multisensory cuing paradigm, relevant visual and irrelevant auditory letter stimuli were presented simultaneously from a central position (Weissman et al., 2004). The fMRI results indicated that conflict from distracting auditory letters with relevant visual letters lead to increased activity in both the ACC and the visual cortex (Weissman et al., 2004), which was interpreted as indicating that attention toward the relevant input is enhanced by boosting activity in the sensory cortex of the task-relevant modality. Our present results are in agreement with this assumption, but also demonstrate that multisensory conflict induces increased activity in the relevant visual cortex even when both the task-relevant and the task-irrelevant components occur at completely different spatial locations.

Auditory cortex

For the auditory cortex, we hypothesized two alternative possible outcomes. One possibility was that attention would spread more to the task-irrelevant auditory stimulation for the congruent than for the incongruent trials because once the incongruency was detected the brain would respond by sending inhibitory signals to suppress auditory cortex processing of the conflicting auditory stimulus. The alternative hypothesis was that the incongruent auditory stimulation would serve as a greater distractor, tending to capture attentional resources and thus lead to greater activity in auditory cortex. Our results clearly provide evidence for the second hypothesis.

The attention-related interaction activity patterns in the auditory cortices differed between the left and right hemispheres. In the right auditory cortex, we found significant 'spread-of-attention' activity during incongruent trials, whereas in the left auditory cortex there were significant, and similar, levels of this activity for both the incongruent and congruent conditions. Thus, put a different way, the 'spread-of-attention' activity was robustly present for conflicting sound in both auditory cortices, but only on the left for congruent trials. This pattern thus suggests that a conflicting incongruent sound (even when task-irrelevant) seems to distract more strongly than a congruent one, apparently leading to greater capture of attention, reflected by an increase of activity in auditory cortex. Using auditory deviants as distracting stimuli in an auditory stream of standards, previous brain imaging studies (Watkins et al., 2007; Sabri et al., 2006; Crottaz-Herbette and Menon, 2006) have suggested that attentional capture of non-speech-related auditory distractors results in enhancement of activity bilaterally in the auditory cortices compared to non-distracting auditory standard stimuli. In addition, a recent EEG study on auditory deviant perception (Jankowiak and Berti, 2007) indicated that deviant stimuli with probabilities as high as 33% are able to induce distraction as shown by an increase of the N2 and P3 ERP components in deviant stimuli occurring in a stream of standards. Accordingly, it is reasonable that the randomized incongruent multisensory events in our study, which occurred with a probability of 33%, would have been able to induce distraction.

As noted above, some spreading of attention due to the extracted auditory stimulus part was also found for the congruent letter/sound stimulus combinations, but only in the left, language-dominant, auditory cortex. In a previous study by Miller and D'Esposito (2005), left Heschl's Gyrus was reported as being specifically involved in the perception of temporally congruent speech stimuli (i.e. voice/mouth). Here, our recent data extend these studies by showing a specific role for the left Heschl's Gyrus in semantically congruent letter/sound combinations, even when the audio component comes from a different spatial location from the task-relevant visual component.

Previous fMRI studies investigating brain activity related to congruent and incongruent speech combinations (e.g., van Atteveldt et al., 2004, 2007; Hocking and Price, 2008; Murase et al., 2008; Noppeney et al., 2008) have almost always presented stimuli from the

same spatial location, typically centrally (but see Fairhall and Macaluso, 2009). In such studies, spatial attention was thus already in place for the location of both the visual and auditory stimuli (i.e., supramodally) and did not need to spread *spatially* to encompass the simultaneously occurring auditory stimulus. In further contrast to these other studies that used central presentation, our introduction of an unattended stimulus side allows the investigation of conflict processing in the absence of attention. Importantly, our results in the auditory cortex (bar graphs Fig. 4B, C central panels) indicate that differences for incongruent versus congruent stimulation could be found only when the visual part was attended, and not when it was unattended. Thus, in our present data, conflict processing was not due simply to a mismatch between two components of a multisensory stimulus, but rather depended fully on attention.

Some differences in the results of studies (with active attention paradigms) of multisensory conflict effects might be explained by which modality is attended (visual/auditory), the task context, and/or the level of the conflict. Focusing on modality effects, a recent study (Noppeney et al., 2008) showed an incongruency effect in the left auditory cortex when the auditory input (words) was attended and the visual part (pictures) was to be ignored. Such results thus suggest that, if the auditory modality is attended, then the processing of incongruency may tend to be stronger in the auditory cortex of the left hemisphere, which is specialized for language processing. Considering task effects, a cueing study by Weissman et al. (2004) did not find any significant incongruency effects in the auditory cortex when the target modality was visual, possibly indicating a less significant auditory capture effect when subjects' are forewarned by a cue. In regards to conflict level, in Sadaghiani et al. (2009), subjects reported the direction of visual moving dots that were accompanied by either congruent or incongruent auditory directional words (left/right/up/down), which resulted in an effect of incongruency in the left rather than the right auditory cortex. One might speculate that, because the speech and motion stimuli in that study came from very different object categories, an incongruent stimulus combination might evoke conflict more at a language-related response level (in contrast, perhaps, to letter–object stimulus-level representational conflict in our study). This language-related response conflict may therefore more heavily engage activity in the auditory cortex of the language-specialized left hemisphere.

In our case, with visual input being attended, and auditory unattended, we infer that attention flows from the attended visual to the auditory modality (as shown in the original Busse et al. study, 2005). We speculate that, if our hypothesis is correct that the increased activity for incongruent trials reflects in part an attentional distraction by the incongruent stimulus, the tendency toward a right-side dominance of this effect could be related to the right hemisphere being more involved in stimulus-driven, bottom-up attentional capture (e.g., Corbetta and Shulman, 2002). Regardless, however, we also note that the overall pattern here (spreading of attention activity larger for incongruent than congruent trials) was similar in the two auditory cortices, although the difference on the left did not reach significance. Accordingly, these left–right incongruency differences may be more a matter of degree, rather than a strict laterality difference.

Conflict-related correlations between brain and behavior and between brain areas

The idea that an incongruent auditory stimulus, even when task-irrelevant and from a different location, could capture attention and result in greater activity in the auditory cortex could also be assessed in the current experiment with behavioral measures. These measures were then used to perform a correlational analysis between behavioral interference (as indexed by the Incongruent-minus-Congruent RT difference) and the incongruency-modulation of the

spreading-of-attention activity in the auditory cortices. A significant positive correlation was found in the right auditory cortex, but not on the left. Such a finding thus mirrors the hemispheric pattern of results described above in which it was found that only the right auditory cortex showed spreading-of-attention activity being significantly modulated by incongruency.

Additionally, the incongruency modulation of the spreading-of-attention activity in the auditory cortices was found to covary significantly across subjects with the analogous incongruency-modulation seen in the ACC. In a previous combined EEG/fMRI study (Crottaz-Herbette and Menon, 2006), using auditory deviants as distractors in a stream of auditory standard stimuli, it was reported that the more activity an irrelevant auditory deviant elicited in auditory cortex the higher the activity in the ACC. Here, our correlations of auditory cortex and ACC are in agreement with these previous studies, but go beyond by showing that these correlations occur also during a multisensory paradigm when the auditory stimuli are task-irrelevant, and even arising from a different location in space.

Our results considered as a whole suggest an overall systems-level model of the cascade of functional activation. With an attentionally-demanding behavioral goal, it is essential to be able to focus on information within the relevant modality. Task-irrelevant conflicting stimuli from the another modality may capture attention away from this goal, which our data suggest is reflected within trial by enhanced activity in the sensory cortex of the task-irrelevant conflicting modality. Such distraction may then be overcome by a combination of ACC activation (Weissman et al., 2004; Egner and Hirsch, 2005, for intravisual sequential effects; and as in our data) and boosting activity to the sensory cortex of the relevant modality (Noppeney et al., 2008; Weissman et al., 2004; Egner and Hirsch, 2005, with sequential effects; and in our data).

Summary and conclusions

In summary, the present study reports several key findings. First, the results confirmed our predictions for the anterior cingulate cortex and visual cortex. In both regions, we observed significant enhancements for incongruent versus congruent letter–sound combinations when the visual component was attended, presumably reflecting the detection of the conflict (in the ACC) for incongruent trials and the need to boost attention to the relevant visual stimulus (in visual cortex) for such trials.

For the auditory cortex, we suggested two competing hypotheses: one which predicted greater 'spreading of attention' for congruent stimuli, as they would assist the visual discrimination and be drawn into the penumbra of attention, and one in which incongruent stimuli would serve as distracters and tend to capture attention and processing resources. The results support the second hypotheses, namely that the spread-of-attention activity was greater if the spatially discordant auditory stimulation was incongruent with the task-relevant visual stimulus. In addition, some spreading-of-attention activity was also found for congruent letter/sound combination, but mainly in the language-dominant left auditory cortex, suggesting some semantic, and thus object-related, dependence of the attentional spreading.

In conclusion, the overall pattern of the present results indicate increased spreading of attention and other attentional interactions for incongruent than for congruent multisensory stimulation in the ACC, the visual cortex, and the auditory cortex. This pattern of activity is consistent with the model in which, as attention spreads from a relevant visual stimulus to a simultaneous but incongruent auditory stimulus, even when it is task-irrelevant and from a different spatial location, the incongruency results in increased distracter-related activity in auditory cortex. This could then lead to enhanced activity of the ACC, presumably related to conflict detection of some sort, followed by enhanced boosting of processing in the visual cortex to

help counter the distracting influence of the auditory modality. Future studies using brain-activity measures with higher temporal resolution (such as EEG) will be important for being able to delineate the precise temporal cascade of brain activation during the 'spreading of attention' for congruent and incongruent multisensory stimulation.

Acknowledgments

References

Abrams, R.A., Law, M.B., 2000. Object-based visual attention with endogenous orienting. Percept. Psychophys. 62, 818–833.

Billington, J., Baron-Cohen, S., Bor, D., 2008. Systemizing influences attentional processes during the Navon task: an fMRI study. Neuropsychologia 46, 511–520.

Burock, M.A., Buckner, R.L., Woldorff, M.G., Rosen, B.R., Dale, A.M., 1998. Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. NeuroReport 9, 3735–3739.

Busse, L., Roberts, K.C., Crist, R.E., Weissman, D.H., Woldorff, M.G., 2005. The spread of attention across modalities and space in a multisensory object. Proc. Natl Acad. Sci. USA 102, 18751–18756.

Christensen, M.S., Ramsøy, T.Z., Lund, T.E., Madsen, K.H., Rowe, J.B., 2006. An fMRI study of the neural correlates of graded visual perception. Neuroimage 31, 1711–1725.

Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. 3, 201–215.

Crottaz-Herbette, S., Menon, V., 2006. Where and when the anterior cingulate cortex modulates attentional response: combined fMRI and ERP evidence. J. Cogn. Neurosci. 18, 766–780.

Dalton, P., Spence, C., 2007. Attentional capture in serial audiovisual search tasks. Percept. Psychophys. 69, 422–438.

Dorsaint-Pierre, R., Penhune, V.B., Watkins, K.E., Neelin, P., Lerch, J.P., Bouffard, M., Zatorre, R.J., 2006. Asymmetries of the planum temporale and Heschl's gyrus: relationship to language lateralization. Brain 129, 1164–1176.

Egly, R., Driver, J., Rafal, R.D., 1994. Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. J. Exp. Psychol. Gen. 123, 161–177.

Egner, T., Hirsch, J., 2005. The neural correlates and functional integration of cognitive control in a Stroop task. Neuroimage 24, 539–547.

Fairhall, S.L., Macaluso, E., 2009. Spatial attention can modulate audiovisual integration at multiple cortical and subcortical sites. Eur. J. NeuroSci. 29, 1247–1257.

Fan, J., Flombaum, J.I., McCandliss, B.D., Thomas, K.M., Posner, M.I., 2003. Cognitive and brain consequences of conflict. Neuroimage 18, 42–57.

Fan, J., Hof, P.R., Guise, K.G., Fossella, J.A., Posner, M.I., 2008. The functional integration of the anterior cingulate cortex during conflict processing. Cereb. Cortex 18, 796–805.

Fiebelkorn, I.C., Foxe, J.J., Molholm, S., 2010. Dual mechanisms for the cross-sensory spread of attention: how much do learned associations matter? Cereb. Cortex 20, 109–120.

Friston, K.J., Ashburner, J., Frith, C.D., Poline, J.-B., Heather, J.D., Frackowiak, R.S.J., 1995. Spatial registration and normalization of images. Hum. Brain Mapp. 2, 165–189.

Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, D., Turner, R., 1998. Event-related fMRI: characterizing differential responses. NeuroImage 7, 30–40.

Hocking, J., Price, C.J., 2008. The role of the posterior superior temporal sulcus in audiovisual processing. Cereb. Cortex 18, 2439–2449.

Jankowiak, S., Berti, S., 2007. Behavioral and event-related potential distraction effects with regularly occurring auditory deviants. Psychophysiology 44, 79–85.

Macaluso, E., Frith, C.D., Driver, J., 2000. Modulation of human visual cortex by crossmodal spatial attention. Science 289, 1206–1208.

Mazziotta, J.C., Toga, A.W., Evans, A., Fox, P., Lancaster, J., 1995. A Probabilistic atlas of the human brain: theory and rationale for its development. NeuroImage 2:89–101.

Miller, L.M., D'Esposito, M., 2005. Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. J. Neurosci. 25, 5884–5893.

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.

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., Zilles, K., 2001. Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. Neuroimage 13, 684–701.

Müller, N.G., Kleinschmidt, A., 2003. Dynamic interaction of object- and space-based attention in retinotopic visual areas. J. Neurosci. 23, 9812–9816.

Murase, M., Saito, D.N., Kochiyama, T., Tanabe, H.C., Tanaka, S., Harada, T., Aramaki, Y., Honda, M., Sadato, N., 2008. Cross-modal integration during vowel identification in audiovisual speech: a functional magnetic resonance imaging study. Neurosci. Lett. 434, 71–76.

Noesselt, T., Bergmann, D., Hake, M., Heinze, H.J., Fendrich, R., 2008. Sound increases the saliency of visual events. Brain Res. 1220, 157–163.

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.

Penhune, V.B., Zatorre, R.J., MacDonald, J.D., Evans, A.C., 1996. Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurement from magnetic resonance scans. Cereb. Cortex 6, 661–672.

Picard, N., Strick, P.L., 2001. Imaging the premotor areas. Curr. Opin. Neurobiol. 11, 663–672.

Poldrack, R., 2007. Tools of the trade: region of interest analysis for fMRI. SCAN 2, 67–70.

Rademacher, J., Morosan, P., Schormann, T., Schleicher, A., Werner, C., Freund, H.J., Zilles, K., 2001. Probabilistic mapping and volume measurement of human primary auditory cortex. Neuroimage 13, 669–683.

Sabri, M., Liebenthal, E., Waldron, E.J., Medler, D.A., Binder, J.R., 2006. Attentional modulation in the detection of irrelevant deviance: a simultaneous ERP/fMRI study. J. Cogn. Neurosci. 18, 689–700.

Sadaghiani, S., Maier, J.X., Noppeney, U., 2009. Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. J. Neurosci. 29, 6490–6499.

Tzourio, N., Nkanga-Ngila, B., Mazoyer, B., 1998. Left planum temporale surface correlates with functional dominance during story listening. NeuroReport 9, 829–833.

Uncapher, M.R., Rugg, M.D., 2005. Effects of divided attention on fMRI correlates of memory encoding. J. Cogn. Neurosci. 17, 1923–1935.

van Atteveldt, N., Formisano, E., Goebel, R., Blomert, L., 2004. Integration of letters and speech sounds in the human brain. Neuron 43, 271–282.

van Atteveldt, N.M., Formisano, E., Goebel, R., Blomert, L., 2007. Top-down task effects overrule automatic multisensory responses to letter–sound pairs in auditory association cortex. Neuroimage 36, 1345–1360.

van Veen, V., Carter, C.S., 2005. Separating semantic conflict and response conflict in the Stroop task: a functional MRI study. Neuroimage 27, 497–504.

Wall, M.B., Walker, R., Smith, A.T., 2009. Functional imaging of the human superior colliculus: an optimised approach. Neuroimage 47, 1620–1627.

Watkins, S., Dalton, P., Lavie, N., Rees, G., 2007. Brain mechanisms mediating auditory attentional capture in humans. Cereb. Cortex 17, 1694–1700.

Weissman, D.H., Giesbrecht, B., Song, A.W., Mangun, G.R., Woldorff, M.G., 2003. Conflict monitoring in the human anterior cingulate cortex during selective attention to global and local object features. Neuroimage 19, 1361–1368.

Weissman, D.H., Warner, L.M., Woldorff, M.G., 2004. The neural mechanisms for minimizing cross-modal distraction. J. Neurosci. 24, 10941–10949.

Weissman, D.H., Warner, L.M., Woldorff, M.G., 2009. Momentary reductions of attention permit greater processing of irrelevant stimuli. Neuroimage 48, 609–615.

Zimmer, U., Macaluso, E., 2007. Processing of multisensory spatial congruency can be dissociated from working memory and visuo-spatial attention. Eur. J. Neurosci. 26, 1681–1691.