Neural Correlates of Coherent Audiovisual Motion Perception

Oliver Baumann[1,2] and Mark W. Greenlee[1]

[1]Institute for Experimental Psychology, University of Regensburg, 93053 Regensburg, Germany and [2]Department of Psychology, University of Oslo, 0317 Oslo, Norway

Real-life moving objects are often detected by multisensory cues. We investigated the cortical activity associated with coherent visual motion perception in the presence of a stationary or moving auditory noise source using functional magnetic resonance imaging. Twelve subjects judged episodes of 5-s random-dot motion containing either no (0%) or abundant (16%) coherent direction information. Auditory noise was presented with the displayed visual motion that was moving in phase, was moving out-of-phase, or was stationary. Subjects judged whether visual coherent motion was present, and if so, whether the auditory noise source was moving in phase, was moving out-of-phase, or was not moving. Performance was greatest for a moving sound source that was in phase with the visual coherent dot motion compared with when it was in antiphase. A random-effects analysis revealed that auditory motion activated extended regions in both cerebral hemispheres in the superior temporal gyrus (STG), with a right-hemispheric preponderance. Combined audiovisual motion led to activation clusters in the STG, the supramarginal gyrus, the superior parietal lobule, and the cerebellum. The size of the activated regions was substantially larger than that evoked by either visual or auditory motion alone. The congruent audiovisual motion evoked the most extensive activation pattern, exhibiting several exclusively activated subregions.

Keywords: audiovisual integration, brain imaging, coherence, multimodal, superior temporal gyrus, supramarginal gyrus

Introduction

Most events in everyday life are perceived simultaneously by different sensory systems. Therefore, the processing and integration of multisensory information are essential for complete perception of our environment and for the planning and control of movements (Calvert and others 2004). With our 2 most important senses, vision and audition, we can perceive the speed and the direction of moving objects. If visual and auditory stimuli are perceived as coincident in space and time, thereby giving the impression that they come from the same source, the information presumably merges, producing a unified percept of movement. The synthesis of sensory information in the brain can contribute to the subject's ability to detect, localize, and discriminate between stimuli, thereby leading to a faster and more precise response (e.g., Miller 1982). The multisensory contribution to motion perception is most pronounced when moving stimuli, encoded by different sensory modalities, occur at the same spatial location and approximately the same time. On the other hand, when the stimuli are presented out of synchrony or from different spatial locations, multisensory enhancement declines (Soto-Faraco and others 2004).

Multimodal stimuli can also attenuate response behavior, for instance, when one of the stimulus components acts as a distractor. One prominent mechanism is multisensory capture (Soto-Faraco and others 2002; Morein-Zamir and others 2003), the phenomenon that an irrelevant stimulus alters the perception of an attended stimulus in a way that leads to a decrease in performance in discrimination or detection tasks.

At the neural level, the integration of visual and auditory information has been investigated foremost with electrophysiological and lesion techniques. Results of electrophysiological (Colby and others 1996; Andersen 1997) and histochemical studies (Rizzolatti and others 1997) in macaque monkeys indicate that the posterior parietal cortex and the premotor cortex are important for the integration of neural signals from different modalities, as well as for the control of movements guided by visual, auditory, and tactile stimuli. Anatomical data also show that the ventral intraparietal area gets direct input from primary visual and auditory areas (Lewis and Van Essen 2000).

Using functional magnetic resonance imaging (fMRI) Bremmer and others (2001) showed that moving audiovisual stimuli activate homolog areas in the human brain: The posterior parietal cortex in both hemispheres, the right ventral premotor cortex, and the lateral inferior postcentral cortex. Lewis and others (2000) identified the intraparietal sulcus (IPS), the anterior middle fissure, and the anterior insula as regions that play an important role in multisensory integration. Bushara and others (1999) found increased activity in the inferior parietal lobule (IPL) in tasks concerning the integration of spatial information from several modalities.

Inhibitory interactions have also been reported. Using fMRI, Calvert and others (2000) showed that activity in the insula and in the colliculus superior (CS) significantly increases during the presentation of temporally synchronous auditory and visual stimuli, whereas with the presentation of asynchronous auditory and visual stimuli the activity in the insula and the CS decreased. It remains to be determined whether these effects are task specific or due to uncontrolled cognitive or attentive factors (Haxby and others 1994; Shulman and others 1997; Binder and others 1999). A recent study by Kayser and others (2005) supports the idea that multisensory integration can take place even without attention and without feedback from higher cortical areas, by showing supra-additive integration of sound and touch in the auditory cortex in anaesthetized monkeys. However, Fujisaki and Nishida (2005) recently showed that the conscious detection of audiovisual synchrony is slow and postattentive requiring feature tracking.

Functional brain imaging can be combined with psychophysical methods to further contribute to our understanding of multisensory integration. In particular we asked whether regions are active in the human brain in response to audiovisual

motion parallel to those found with electrophysiological recordings in primates.

We investigated the cortical activations associated with coherent visual motion perception in the presence of a stationary or moving sound source. In an fMRI paradigm, our subjects were presented 5-s episodes of random-dot motion containing either no (0%) or abundant (16%) coherent direction information. Simultaneous auditory noise was presented with an in phase moving (with respect to the visual motion), antiphase moving, or stationary sound source. To assure that the subjects attended to both the visual and the auditory stimuli, a 4 alternative forced-choice response paradigm was employed. Subjects had to judge whether visual coherent motion was present, and if so, whether the auditory sound source was moving in phase, was moving out-of-phase, or was not moving.

Using an event-related design, blood oxygen level–dependent (BOLD) responses for trials with congruent (in phase) audiovisual coherent motion were compared with those found for trials with incongruent (antiphase) audiovisual motion and with those on trials containing visual motion and a stationary sound source. Using this approach, we isolated different processes involved in audiovisual motion integration. In contrast to previous studies investigating neuronal responses to audiovisual stimuli (e.g., Lewis and others 2000), our study did not aim to reveal the differences between unimodal and multimodal stimulus conditions, but rather to show the differences between congruent and incongruent stimulus conditions. By doing so we seek to identify brain regions that are involved in multisensory integration while avoiding contamination of the results through common activity (i.e., that related to attention or anticipation of the stimulus), which could result from simply contrasting unimodal and bimodal conditions (Gondan and Roder 2006).

Prior to the fMRI study we conducted a psychophysical study with an independent sample of 11 subjects to determine the effect of the relative phase of the virtual sound source (in phase or antiphase) on the subject's ability to detect visual coherent motion. Here we used 4 levels of motion coherence to evaluate the effect of the sound source on the detection of visual coherence motion around the subject's threshold.

Materials and Methods

Psychophysical Study

Subjects

The psychophysical study was conducted outside of the scanner with an independent group of 11 right-handed healthy volunteers (8 female). All subjects gave informed consent to procedures approved by the Regensburg University's ethics committee. The subjects' ages ranged from 19 to 28 years (mean age, 22 years). The subjects were assessed in a training session to exclude subjects who were unable to identify at least 80% of all "hit"-trials the sound direction (in phase or antiphase) correctly, those with false-alarm rates >20%, and those who could not maintain stable fixation. Three subjects were excluded in the training session for one or more of these reasons.

Visual Stimulation

The subjects rested their head on a chinrest and viewed the stimuli on a Sony cathode ray tube (CRT) monitor. The stimuli were presented using a Matrox graphic card. The image was 26° of visual angle horizontal and 20° of visual angle vertical (1024 × 768 pixels) at a viewing distance of 70 cm.

The stimuli were digital movies created with Matlab (Version 6.5). The visual stimuli consisted of a white fixation target dot (0.4°, 100 cd/m^2) and 400 sparse gray background dots (0.4°, 45 cd/m^2) on a black background (0.5 cd/m^2). The white stationary fixation dot was displayed in the center of the display. The background dots moved along random trajectories creating a random-dot kinematogram (RDK). Four levels of motion coherence (0%, 4%, 6%, and 8%) were presented evenly distributed using the method of constant stimuli. In the conditions with coherent motion, the coherent dots moved along the horizontal axis with a sinusoidal velocity profile. The maximum speed of 12.6°/s occurred when the dots passed the center of the screen. This yielded a frequency of 0.2 Hz. The speed of the random-dot trajectories was distributed over the same range and had the same mean velocity as the coherent dots. The half-life of each dot (coherent or random) was 1 s, after which it was replaced by another dot with a new speed and direction. These transition periods were randomized over time, such that a steady migration of dots from random to coherent or vice versa occurred.

Auditory Stimulation

The moving sound was Gaussian white noise which was convolved with generic head-related transfer function for positions +/− 12° of azimuth angle, in discrete steps of 1°. The sounds generated were smoothed by a hanning window to create the impression of a smoothly moving sound source. The virtual sound source had the same sinusoidal velocity profile as the coherently moving dot.

The stationary sound was Gaussian noise, which was convolved with the same generic head-related transfer function for the position 0° of azimuth angle (i.e., straight ahead). This manipulation yielded the impression of a stationary sound source located just in front of the listener. The moving and the stationary sound files had the same mean energetic profile. The amplitude was ~76 dB(A) sound pressure level (SPL) maximum inside the headphones. In the psychophysical study, the acoustic noise was presented using a Soundblaster soundcard, a digital amplifier and Beyer dynamic DT 990 headphones.

Audiovisual Stimulation

The visual and auditory stimuli were merged together using an audiovisual editing program (FX RESound, Hepple, Inc., Hewitt, Texas), leading to 5-s episodes of audiovisual digital movies. Overall, 7 combinations of stimuli were constructed containing either no (0%) or 1 of 3 levels (4%, 6%, 8%) of coherent direction information and either an in phase moving or antiphase moving sound source. The sequence of trials from the different conditions was randomized and the direction of the auditory stimulus was counterbalanced for all conditions. In the condition with 0% visual coherence, the sound source was designated as moving because phase relative to the coherent motion cannot be defined. The stimuli were presented using "Presentation" Version 9.20 (Neurobehavioral Systems, Inc., Burnaby, BC, Canada).

Task

Subjects judged whether visual coherent motion was present in the random-dot displays, and if so, whether the auditory sound source was moving in phase or was moving in antiphase. They were instructed to press 1 of 3 different buttons depending on whether they thought the stimulus contained any dots that moved in a coherent direction and the auditory sound source was moving in phase, in antiphase, was not moving, or the stimulus contained no dots that move in a coherent direction, regardless of the auditory condition. A high tone signaled the subjects that they had judged correctly and a low tone that they had made a wrong decision. The accuracy of the responses given by the subjects was estimated by d'. From the collected hits and false alarms, the sensitivity measure d' can be computed as follows.

$$d' = Z_{SN} - Z_N = Z(P(\text{hits})) - Z_N = Z(P(\text{false alarms})).$$

The value of d' corresponds to the distance between 2 standard normal distributions that model the noise associated with discriminating a signal from no signal. A d' of 0 denotes that the subject cannot reliably detect the signal. The larger the d', the higher is the detectability of the signal. Thus, d' is a measure of detectability, which is independent of the response criterion of the subject. A detailed description of this procedure can be found in Wickens (2001).

For computational purposes, stimuli that had a visual coherence level of 0% were designated as noise trials and stimuli with a visual coherence

level of 4%, 6% and 8% were designated as signal trials. Pressing buttons 1 and 2 were treated as "yes" responses (signal present) and pressing button 3 as "no" (noise only), respectively. In total, every subject conducted 320 trials, 80 for each level of visual coherence (with equal proportion of audiovisual in phase and antiphase conditions).

On trials in which coherent dots were present and the subject pressed button 1 or 2 the response was counted as a hit. Therefore, the estimation of d' was not influenced by whether subjects were correct with respect to relative direction. For the estimation of the detectability of coherent visual motion, a response was counted as a hit even if the subject misjudged the relative phase of the auditory and the visual stimulus. Subjects were instructed to wait to respond at the end of the stimulation, to be as accurate as possible. Therefore, response times were determined from the offset of the stimulus to activation of the response button. In all trials, subjects responded within the 3-s time window allowed. The subjects were instructed to maintain steady fixation on the fixation dot during the entire experiment. During the rest periods between 2 stimuli a blank screen was presented for the duration of 4 s. In total 320 trials were presented, separated in 2 blocks of duration of 24 minutes each. The 2 experimental blocks were conducted on 2 subsequent days.

Recordings of Eye Movements

During the psychophysical measurement, eye movements were recorded to monitor fixation. Eye movements were recorded using the IRIS-Eyetracker (Skalar, Delft, NL), a limbus tracking device (Reulen and others 1988). The Matlab Data Acquisition Toolbox was used to acquire the signals derived from the IRIS-Eyetracker. The sampling frequency of the eye-tracker signal was 500 Hz, the spatial resolution was 0.1°. The eye-recording system was calibrated with 4 eccentricities (−10°, −5°, +5°, +10°), to determine the deviation from the fixation position. Using the Matlab Signal Processing Toolbox, we analyzed the eye trajectories offline and evaluated the fixation performance of the subjects. In all conditions, the maximum deviations during stimulus presentation were <0.1°. According to this, all of the subjects could maintain a stable fixation during the stimulus presentation. Trials on which the subject broke fixation and initiated pursuit or saccadic tracking would have been eliminated from the analysis but this turned out to be unnecessary.

fMRI-Study

Subjects

An independent group of 12 right-handed volunteers (8 female) participated after giving informed consent to procedures approved by the Regensburg University's ethics committee. The subjects' ages ranged between 18 and 35 years (mean 22 years). None of the subjects had a history of neurological or psychiatric disorders. All subjects had no known hearing or visual impairments. All subjects participated in a training session during which they practiced the audiovisual motion task. Subjects' performance was assessed in the psychophysical laboratory prior to imaging to exclude subjects who could not fulfill the criteria described above.

Visual Stimulation

The subjects were positioned supine in the scanner with their head tightly secured in the headcoil to minimize head movement. They viewed the stimuli with a mirror that reflected the image from the projection screen placed at the head of the subject in the end of the scanner gantry.

The same stimulus sequences were presented as in the psychophysical study. The visual stimuli consisted of a white fixation target dot (0.4°, 250 cd/m²) and 400 sparse gray background dots (0.4°, 110 cd/m²) on a black background (5 cd/m²). In half of the trials, the RDKs contained 16% coherently moving dots and in the other half there was no coherent motion present. We chose the 16% coherence level to guarantee that all subjects could detect the coherent visual motion in most of the trials.

Auditory Stimulation

The parameters of the auditory stimulation were the same as in the psychophysical study. The acoustic noise was presented using a

Soundblaster soundcard, MR Confon amplifier, and MRI-compatible sound-dampening headphones (MR Confon, GmbH, Magdeburg, Germany). The sound pressure level (SPL) of the auditory noise stimuli was 76 dB(A) and as such comparable with those used in the psychophysical study.

Audiovisual Stimulation

Overall, 6 combinations of stimuli were constructed containing either no (0%) or abundant (16%) coherent direction information and either an in phase moving, antiphase moving, or stationary sound source. In the condition with 0% visual coherence, the sound source was designated as moving because phase relative to the coherent motion cannot be defined. The sequence of trials from the different conditions was randomized, and the direction of the auditory stimulus was counterbalanced for all conditions.

Task

The subjects judged whether visual coherent motion was present in the random-dot displays, and if so, whether the auditory sound source was moving in phase, was moving in antiphase, or was not moving. Responses were recorded with a 5-button fiber-optic response box (Lumitouch, Photon Control, Ltd, Burnaby, BC, Canada). Subjects were instructed to press 1 of 4 different buttons depending on whether they thought the stimulus contained any dots that moved in a coherent direction and the auditory sound source was moving in phase, or in antiphase, not moving, or whether the stimulus contained no dots that move in a coherent direction, regardless of the auditory condition. There was no auditory feedback to avoid confounding artifacts with respect to activation in the auditory cortex. The accuracy of the responses given by the subjects was measured in units of d', in analogy to the psychophysical study. Pressing buttons 1, 2, and 3 were treated as "yes" responses (signal present) and pressing button 4 as "no" (noise only), respectively.

Subjects were instructed to wait to the end of the stimulation to respond thereby avoiding confounding artifacts with respect to activation in motor areas. Therefore, response times were measured from the offset of the stimulus to activation of the response button. In all trials, the subjects responded within the 6-s time window allowed.

Subjects were instructed to maintain steady fixation on the bright white fixation dot during the entire experiment. Between 2 stimuli, a static image containing random dots was presented for the duration of 10 s. This image also contained the white fixation dot. This image was presented to prevent dark adaptation during the interstimulus interval and to maintain a steady-state level of stimulation. In total, 120 trials were presented, which required a total duration of 30 min.

Recordings of Eye Movements

Eye movements were recorded using the MR-Eyetracker (CRS, Ltd, Rochester, England), a fiber-optic limbus tracking device (Kimmig and others 1999). The Matlab Data Acquisition Toolbox was used to record the signals derived from the MR-Eyetracker. The sampling frequency of the eye-tracker signal was 1000 Hz; the spatial resolution was 0.1°. The eye-recording system was calibrated with 4 eccentricities (−15, −20, +15, +20°) to determine the deviation of the fixation position.

Using the Matlab Signal Processing Toolbox, we analyzed the resulting eye trajectories offline and evaluated the fixation performance of the subjects. The maximum deviations were in all conditions <1°/s, which was due to baseline drifts and noise. As in the training session outside the scanner, all subjects were able to maintain stable fixation.

MR Imaging

MRI was performed with a 1.5-Tesla clinical scanner (Magnetom Sonata, Siemens, Erlangen, Germany) equipped with an echo-planar imaging (EPI) booster for fast gradient switching and an 8-channel phase array full-head radio-frequency receive–transmit headcoil (MR-Devices). High-resolution, sagittal T1-weighted images were acquired with the magnetization prepared, rapid acquisition gradient echo sequence to obtain a 3D anatomical scan of the head and brain. Functional imaging was performed with T2*-weighted gradient EPI. We used a variation of Hall's sparse temporal sampling technique (Belin and others 1999; Hall and others 1999) to circumvent interference from acoustic noise

created by the gradient coils, such that the onset of the MR acquisition began immediately after the end of the audiovisual stimulation. The acquisition time was 3.3 s, with an adjacent waiting period of 11.7 s, resulting in a total time repetition of 15 s. The time to echo corresponded to time echo = 60 ms, the flip angle corresponded to 90°, and we used a field of view (FOV) = 192 mm, with a voxel matrix of 64×64, resulting in a voxel size of $3 \times 3 \times 3$ mm. We acquired volumes with 36 slices, aligned parallel to the anterior and posterior commissures (AC-PC) line, with a gap of 0.45 mm between slices and could thus image nearly the entire neocortex, with the only exception of the most anterior part of the inferior temporal cortex. The stimulation protocol for a single experimental run consisted of 120 alternating periods of stimulation and rest (stationary visual noise), resulting in a total of 120 volumes per subject.

fMRI Data Analysis

The data were preprocessed and analyzed on single subject level using Statistical Parametric Mapping, version 2 (SPM2). After motion correction, the functional images were coregistered to the anatomical volume to normalize both to the Montreal Neurological Institute (MNI) Template (Friston and others 1995a). Functional images were smoothed with a 3D-Gaussian kernel (full width, half maximum, FWHM = 8 mm).

Analysis using the general linear model (Friston and others 1995b) was done after applying high-pass filtering (cut-off: 128 s). In an epoch design analysis, responses during the 5-s stimulation periods were modeled with a boxcar convolved with the hemodynamic response function separately for the 5 conditions (0% visual coherence with moving auditory noise, 0% visual coherence with stationary auditory noise, 16% visual coherence with in phase auditory noise, 16% visual coherence with antiphase auditory noise, 16% visual coherence with stationary auditory noise). For the random-effects group analysis we used the nonparametric SnPM-Toolbox (Holmes 1994; Holmes and others 1996). For each interesting difference in effect sizes we calculated 1 contrast image per subject representing this difference on an individual level. These images were analyzed on the group level with the SnPM Test for "multiple subjects, 1 scan per subject," the nonparametric equivalent of a t-test. The only assumption this method uses is that the contrast value of nonactivated voxels distribute evenly around zero. A 3D variance smoothing using a FWHM of 8 mm was performed. Variance smoothing can enhance the power of the group analysis even above the parametric methods of Gaussian random fields if the assumption of sufficient smoothness of the parametric maps is violated. For small group sizes this is often the case. Voxels surpassing a statistical threshold of $P = 0.05$ (Tmax-contrast analysis, corrected for multiple comparisons) were identified as activated. MNI coordinates were transformed to Talairach coordinates, which we report here. The transformation was performed with the Wake Forest University-Pickatlas (Lancaster and others 1997, 2000; Maldjian and others 2003). The SPM2 extension MNI Space Utility (MSU) by S. Pakhomov was used for the identification of anatomical locations. This tool relies on the mni2tal program combined with data of the Talairach demon (Lancaster and others 2000). The functional group data were mapped to the Human Colin surface-based atlas (Van Essen and others 2001) with the Caret Map fMRI to Surface computer program (Van Essen 2002).

Results

Psychophysical Study

For the majority of the subjects, d' was higher in the in phase condition than in the antiphase condition (for 9 of 11 subjects in the 4% condition, for 5 of 8 in the 6% condition, and for 8 of 11 in the 8% condition). A Wilcoxon test revealed, that the difference between the in phase and antiphase conditions (averaged over all visual coherence levels) is highly significant ($P < 0.001$). For all subjects, d' increased with the visual coherence level (Fig. 1a). Additionally, the d' values of all 3 visual coherence levels (averaged over the sound conditions) were significantly different from each other (Friedman, $P < 0.001$) and the difference between the 2 sound conditions within the 4% (Wilcoxon, $P = 0.013$) and 8% (Wilcoxon $P = 0.028$) visual coherence level was also significant.

The response times were in general longer in the antiphase condition, but the difference between the 2 auditory conditions was not statistically significant (Wilcoxon, $P = 0.131$). Likewise, there were no significant differences between the different levels of visual coherence (Friedman, $P = 0.797$) (Fig. 1b). Because subjects were instructed to wait until the end of the stimulation to respond, these results were not unexpected.

fMRI Study

In the fMRI study, most of the subjects had a hit rate near 100% and all the 12 subjects had a false-alarm rate below 20%. The mean d' value was 3.92 (standard error [SE] 0.17) for the in phase condition, 3.73 (SE 0.75) for the antiphase condition, and 4.03 (SE 0.64) for the stationary sound condition. A Friedman test revealed that the d' of the 3 conditions were not significantly different from each other ($P = 0.227$). Owing to the ceiling effect evoked by the relatively high coherence level of 16% it is not surprising that a significant difference was not evident.

The response times were not significantly different in the 3 sound conditions (Friedman, $P = 0.424$), and there was no significant (Wilcoxon, $P = 0.433$) difference in response times between the 2 levels of visual coherence (i.e., 0% and 16%). The average response time in all conditions was 837 ms (SE 65 ms). Because subjects were instructed to wait until the end of the stimulation to respond, these results were not unexpected.
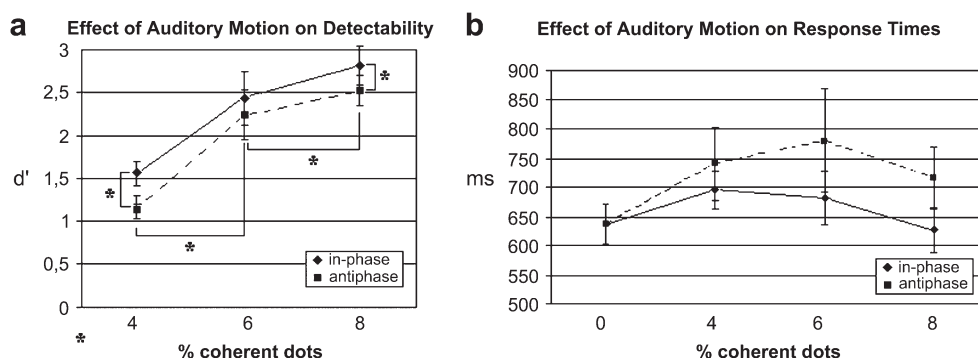
Figure 1. (a) Mean d' values and standard errors of 11 subjects for the in phase and antiphase condition as a function of motion coherence. (b) Mean response times (measured from the offset of the stimulus) and standard errors of eleven subjects for the same experimental conditions.

Functional MRI Data

Moving versus Static Auditory Stimuli with Random-Dot Motion

The results of the across-subjects analysis for the condition with 0% coherent dots and moving sound with respect to the condition with 0% coherent dots and stationary sound revealed 2 activation clusters. The BOLD clusters were located in the superior temporal gyrus (STG) (area 42) in both hemispheres corresponding to the secondary auditory cortex. The left-hemispheric activation cluster extended also to the supra-marginal gyrus (SMG) (area 43). The right-hemispheric cluster was thereby about a factor 3 larger than the left-hemispheric one (Table 1, Fig. 2*a,b*). These regions of activations are in close agreement with the study of Baumgart and others (1999), who reported an area in associative auditory cortex that responded selectively to a moving sound source.

Combined Audiovisual Motion: Moving versus Static Acoustic Noise in Presence of Coherent Visual Motion

Modulation of the neural activity during the perception of an abundant (16%) coherent visual motion stimulus by the in phase and antiphase auditory noise conditions is shown in Figure 2(*a,b*).

The comparison between the conditions in phase versus the static sound source led to 2 large (right hemisphere: 433 voxels, left hemisphere: 354 voxels) significant activation clusters, which were located in the STG (area 22 and 42) and the SMG (area 43) in both hemispheres. Another small left-hemispheric cluster was active in the precuneus (area 5) (Table 1, Fig. 2*a,b*).

The comparison between the conditions antiphase and the static sound source led to bihemispheric activations located in the STG (area 22 and 42) and the SMG (area 43) (Table 1, Fig. 2*a,b*). As can be seen in Figure 2(*a,b*), the activation clusters are located in the same brain regions as in the in phase condition but the activations are smaller by about a factor of 4.

Detection of Coherent Visual Motion: Coherent Visual Motion versus Random Visual Motion

Our experimental design allowed us to isolate the effect of coherent visual motion on brain activation in the presence of

stationary sound. The results of the SnPM across-subject, random-effects analysis for the condition with 16% coherent dots and stationary sound with respect to the condition with 0% coherent dots and stationary sound revealed 2 right- and 2 left-hemispheric activation clusters with a spatial extend ranging between 33 and 47 voxels per cluster. The right-hemispheric activation clusters were located in the superior parietal lobule (areas 7 and 5) and in the SMG extending to the STG (area 40 and 22). The first left-hemispheric activation was located in the precentral gyrus (area 6) and the second in the precentral gyrus near to the middle frontal gyrus (areas 6 and 4) (Table 2, Fig. 3*a,b*).

Effect of the Moving Sound Source on the Response to Coherent Visual Motion

To investigate the effect of a moving sound source on neural activation, we compared the activations evoked when the moving auditory noise was in phase with the visual coherent motion and when the sound source was in antiphase. In the condition with 0% visual coherence, the sound source was designated as moving because phase relative to the coherent motion cannot be defined.

The contrast "moving acoustic noise and 16% in phase coherent visual motion greater than 0% coherent visual motion" led to several extended activated brain regions (in total 6477 activated voxels), in the superior parietal lobule, SMG, lingual gyrus in both hemispheres, and in the cerebellum. Other activated regions were located in the right hemisphere: in the superior frontal gyrus (SFG), the cuneus, the cingulate gyrus, the STG, and the caudate nucleus. The cuneus, the middle frontal gyrus, and the inferior frontal gyrus were active only in the left hemisphere (Table 2, Fig. 3*a,b*).

The contrast between the condition moving acoustic noise with 16% antiphase coherent visual motion and the condition with 0% coherent motion revealed several activated regions in both hemispheres in the superior parietal lobule, the SFG, the SMG, the STG, the cuneus, and the precentral gyrus. The cerebellum was only activated on the right side. The middle frontal gyrus and the lingual gyrus were exclusively activated in the left hemisphere (Table 2, Fig. 3*a,b*). Compared with the in phase condition, the activation patterns of the antiphase sound condition were less pronounced (in total 1241 activated voxels). This observation is especially true for the activated regions in the superior, the intraparietal area, the SMG, the STG, and the SFG (Table 2, Fig. 3*a,b*).

Discussion

We examined the facilitating and inhibitory effects of an in phase, antiphase, or stationary sound source on the perception of coherent visual motion dot stimuli, as well as its effects on the resultant brain activations. We asked the following questions:

1. Does in phase auditory noise facilitate the detection of coherent visual motion?
2. How is the neuronal activity during the perception of visual coherent motion influenced by in phase, antiphase or stationary auditory stimuli, respectively?

We first discuss the pattern of BOLD responses found here and compare them with earlier results. Subsequently, we attempt to describe how these results reveal the extent to

Table 1

Talairach coordinates (*x*, *y*, and *z*) of the maximum pseudo *t*-value within each cluster for the investigated contrasts

Region	Hemisphere	Brodmann area	Talairach coordinates			Pseudo *t*-values of maxima (clustersize in number of voxels)
			x	*y*	*z*	
0% coherent visual motion: moving acoustic noise > stationary acoustic noise						
STG	R	42	51	−25	12	6.62 (108)
STG/SMG	L	42/43	−50	−27	11	6.07 (39)
16% coherent visual motion: in phase acoustic noise > stationary acoustic noise						
STG/SMG	R	22/42/43/40	50	−23	10	7.60 (433)
STG/SMG	L	47/42/43	−53	−25	10	7.38 (354)
Prec	L	22	−8	−50	56	6.13 (48)
16% coherent visual motion: antiphase acoustic noise > stationary acoustic noise						
STG/SMG	L	40/43	−55	−25	12	6.28 (87)
SMG	R	22	63	−38	13	6.47 (69)
STG	R	22	51	−21	8	6.36 (38)
STG	R	6/4	55	−24	16	5.45 (10)

Note: For each cluster the hemisphere, Brodmann areas, and anatomical structures are specified, in which the respective cluster is located. Significant clusters of at least 10 contiguous voxels or more with a statistical threshold of alpha = 0.05 (corrected for multiple comparisons) are presented. Prec = precuneus.
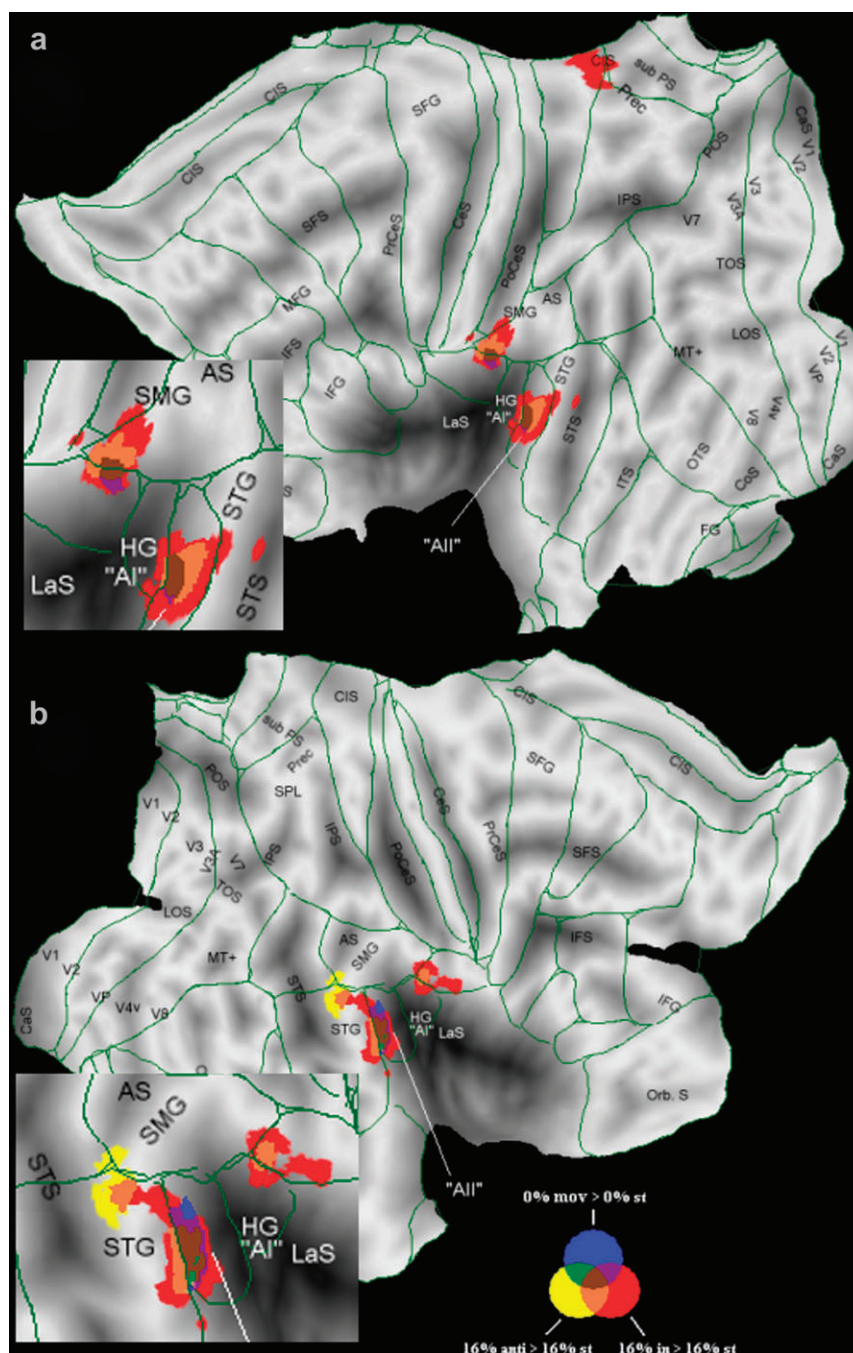
Figure 2. (a) Left-hemispheric group activation maps for moving versus stationary sounds of 12 subjects with detail magnification of the activated region. (b) Right-hemispheric group activation maps for moving versus stationary sounds of 12 subjects with detail magnification of the activated region. The different colors indicate the effect of the visual RDK on this activation (red = 16% in phase moving > stationary, yellow = 16% antiphase moving > stationary, blue = 0% moving > stationary). Flat map representation of significant fMRI activity. Overlaps are indicated by intermediate colors (see color inset). Activation is shown overlaid onto MNI-normalized single subject right hemisphere flat map (Van Essen 2002) template (significant clusters surpassing a threshold of alpha = 0.05 [corrected for multiple comparisons] are presented). Identified visual areas (V1, V2, MT+, etc.) are from the Colin atlas database. The borders represent Brodmann areas from the Colin atlas. *Abbreviations*: AI = primary auditory cortex, AII = secondary auditory cortex, AS = angular sulcus, CaS = calcarine sulcus, CeS = central sulcus, CiG = cingulate gyrus, CiS = cingulate sulcus, CoS = collateral sulcus, FG = fusiform gyrus, GL = lingual gyrus, IFG = inferior frontal gyrus, HG = Heschl's gyrus, ITG = inferior temporal gyrus, ITS = inferiotemporal sulcus, LaS = lateral sulcus, LOS = lateral occipital sulcus, MFG = middle frontal gyrus, Orb. S = orbital sulcus, PoCeG = posterior central gyrus, PoCeS = posterior central Sulcus, Prec = precuneus, PrCeG = precentral gyrus, SFS = superior frontal sulcus, SPL = superior parietal lobule, STS = superior temporal sulcus, subPS = subparietal sulcus, TOS = transverse occipital sulcus.

which a select group of cortical areas underlie our ability to integrate audiovisual motion cues.

Auditory Motion versus Static Auditory Stimuli

With respect to the comparison between moving and stationary sound sources our results are in line with the study of Baumgart and others (1999) who found an extensive cluster of activation mainly in the right planum temporale. In our study, the auditory-motion condition activated extended regions in both hemispheres in the STG (Brodmann area 42), with a right-hemispheric preponderance (Fig. 2a). In agreement with Baumgart and others (1999), our results imply that either the

Table 2

Talairach coordinates (x, y, and z) of the maximum pseudo t-value within each cluster for the investigated contrasts

Region	Hemisphere	Brodmann area	Talairach coordinates			Pseudo t-values of maxima (clustersize in number of voxels)
			x	y	z	
Stationary acoustic noise: 16% coherent visual motion > 0% coherent visual motion						
SPL	R	5/7	40	−44	60	6.04 (47)
PrCeG	L	6	−53	4	40	5.94 (47)
SMG/STG	R	40/22	57	−40	13	5.75 (44)
PrCeG	L	6/4	−36	−9	61	5.84 (33)
Moving acoustic noise: 16% in phase coherent visual motion > 0% coherent visual motion						
SPL/IPS	R + L	5/7	−8	−63	62	7.91 (1699)
SMG/SPL/IPS	L	40/5/7	−63	−26	22	8.45 (1516)
GL/cerebellum	R + L	18/19	−4	−72	−10	7.39 (890)
SFG	R + L	6/4	0	−6	60	7.20 (803)
GL/cuneus	R + L	18/19	8	−70	13	7.42 (410)
CiG	R + L	24	2	−2	42	6.59 (201)
SMG/STG	R	40/22	67	−28	18	6.39 (102)
Nuc. Cau.	R	—	12	10	0	6.35 (88)
MFG/IFG	L	10/46	−43	43	14	6.45 (53)
Cuneus	R + L	19	2	−82	24	5.89 (61)
IFG	R	9	51	9	25	5.97 (59)
Cerebellum	R	—	34	−60	−26	6.10 (53)
PrCeG	L	4	−34	−12	63	5.93 (50)
Nuc. Lent.	L	—	−24	4	6	6.31 (47)
Cuneus	L	19	−18	−84	34	5.77 (43)
IFG	L	9	−53	3	18	5.94 (31)
MTG	L	37	−53	−58	5	5.75 (31)
PrCeG	R	4	56	−4	40	6.19 (29)
Cerebellum	R	—	28	−68	−22	5.48 (25)
CiG	R	24	16	−36	44	5.74 (21)
Thal	L	—	−12	−18	8	6.33 (20)
STG	L	22	−55	−25	5	5.56 (14)
Nuc. Cau.	R	—	18	1	13	5.69 (13)
SPL	L	7	−18	−51	71	5.54 (13)
SPL	L	7	−18	−44	61	5.75 (11)
IFG	L	9	−40	16	26	5.57 (11)
PrCeG	L	—	−60	−21	42	5.52 (10)
Moving acoustic noise: 16% antiphase coherent visual motion > 0% coherent visual motion						
SPL	L	7	−10	−72	55	6.78 (234)
SPL	R	7	20	−67	55	6.16 (139)
SFG	R + L	6	0	−4	62	6.51 (109)
PrCeG	L	4	−40	1	26	6.53 (80)
Cuneus	R + L	18	10	−67	14	6.28 (80)
Cerebellum	R	—	34	−58	−21	6.45 (77)
SPL	R	7	30	−46	48	5.82 (76)
SPL	L	5/7	−32	−44	45	6.09 (65)
STG/SMG	L	22	−61	−40	17	6.09 (50)
MFG	L	10	−38	51	12	6.29 (44)
GL	L	18	−12	−70	−7	5.77 (38)
PrCeG	R	4	20	−13	60	6.17 (36)
STG/SMG	R	22	60	−40	17	5.85 (30)
Cerebellum	R	—	6	−80	−14	5.80 (22)
IPL/IPS	L	40	−61	−35	29	5.59 (48)
Prec	R + L	7	4	−57	62	5.60 (32)
Cuneus	L	19	−18	−84	28	5.67 (17)
Cerebellum	L	—	−26	−72	−12	5.57 (16)
IFG	R	9	42	9	25	5.60 (14)
Nuc. Lent.	R	—	26	6	2	5.67 (13)
Cuneus	R + L	19	0	−80	32	5.51 (11)
ITG	R	19	54	−62	−6	5.52 (10)

Note: For each cluster the hemisphere, Brodmann areas, and anatomical structures are specified, in which the respective cluster is located. Significant clusters of at least 10 contiguous voxels or more with a statistical threshold of alpha = 0.05 (corrected for multiple comparisons) are presented. CiG = cingulate gyrus, GL = lingual gyrus, IFG = inferior frontal gyrus, ITG = inferior temporal gyrus, MFG = middle frontal gyrus, Mtg = middle temporal gyrus, Nuc. Cau. = caudate nucleus, Nuc. Lent. = lentiform nucleus, PoCeG = posterior central gyrus, Prec = precuneus, PrCeG = precentral gyrus, SPL = superior parietal lobule, Thal = Thalamus.

STG contains neurons that are movement selective or this area contains a map of auditory space (see also, Wagner and others 1997; Griffiths and others 1998; Pavani and others 2002). Unlike previous auditory motion studies, which mostly reported that only the right hemisphere is involved, we found left-hemispheric activation, as well.

Interestingly, the auditory region activated in our experiment was also found to be involved during pitch discrimination tasks (Lewis and others 2000; Patterson and others 2004), which indicates that this area is probably not exclusively concerned with motion processing. This conclusion is also corroborated by single-unit studies in monkeys (Ahissar and others 1992).

Audiovisual Motion Perception

The results of the psychophysical study conducted outside the scanner yielded clear evidence that an in phase moving sound source leads to significantly better performance in detecting coherent visual motion, compared with a condition with a sound source moving in antiphase. This corresponds to some extent with the results of Soto-Faraco and others (2004) who asked subjects to judge the direction of apparent auditory motion while they ignored visual apparent motion. If the 2 motion streams were congruent, the subjects correctly judged the direction on almost all trials, whereas in the incongruent condition they performed at chance level. In contrast to our results, Soto-Faraco and others (2004) found the visual stream to be unaffected by the auditory motion and attributed this cross-modal asymmetry to visual capture. However, the fact that auditory motion is not able to capture visual motion does not exclude that auditory motion may influence the detectability for visual motion, as clearly shown in our psychophysical results (Fig. 1).

Brain Regions Subserving Coherent Audiovisual Perception

The results of the SnPM analysis clearly support the role of a select group of cortical areas that underlie audiovisual motion perception. Our evidence suggests the existence of cortical foci, namely the SPL, SMG, IPS, STG, and also areas in the superior frontal and visual cortex that are sensitive to audiovisual information (i.e., in phase audiovisual motion stimuli). The same correspondence was found in the responses of single neurons to multisensory stimuli in monkeys (Stein and Meredith 1993). Using imaging techniques and single-cell recordings, it has been shown that the superior parietal and superior temporal cortex are critically involved in integration of auditory and visual information in human and nonhuman primates (Benevento and others 1977; Hyvarinen and Shelepin 1979; Bruce and others 1981; Rizolatti and others 1981; Lewis and others 2000). Moreover, all these areas were also found to be activated by more than one modality (Bremmer and others 2001; Kayser and others 2005).

As outlined in the results section, we also found BOLD clusters in the STG, SMG, IPS, SPL, and also in striate and extrastriate visual areas, which responded either exclusively or at least more robustly to congruent audiovisual motion stimuli. It appears that the congruent visual information leads to an enhancement of the activity in the auditory association cortex (Fig. 2a,b) and congruent auditory motion leads to integration processes in the higher association areas like the SMG, the superior parietal lobule and the superior frontal cortex (Fig. 3a,b). These findings cannot be explained by an attentional bias toward one sensory modality because attention to both information sources (visual, auditory) was required to perform the task. The activation related to congruent audiovisual stimuli in these areas appears to contribute to our unified perception of multisensory object motion perception. Our results therefore
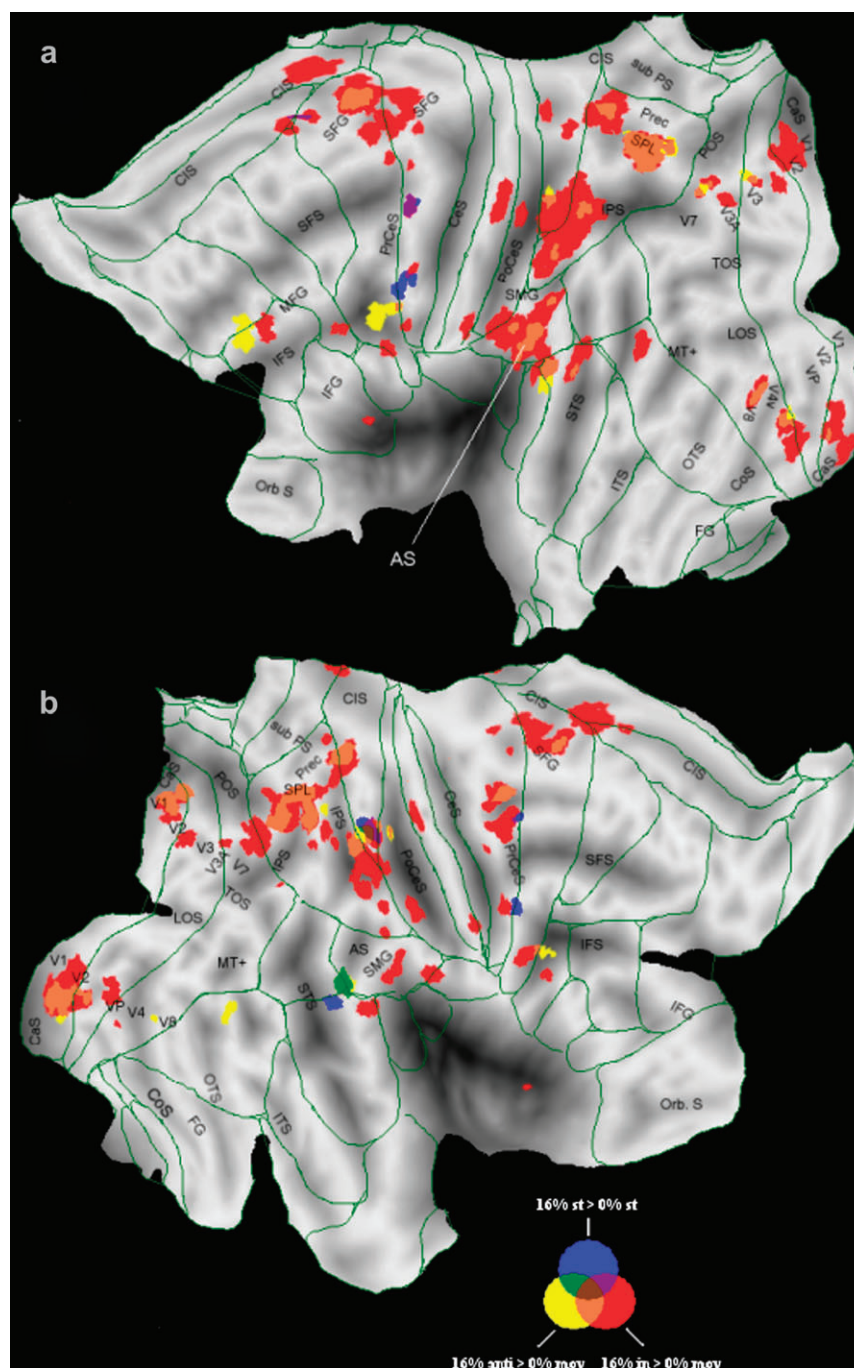
Figure 3. (*a*) Left-hemispheric group activation maps for coherent versus random visual motion of 12 subjects. (*b*) Right-hemispheric group activation maps for the data of 12 subjects. The different colors indicate the effect of the different sound conditions on this activation (red = in phase, yellow = antiphase and blue = stationary). Flat map representation of fMRI activity. Overlaps are indicated by intermediate colors (see color inset). Activation is shown overlaid onto MNI-normalized single subject right hemisphere flat map (Van Essen 2002) template (significant clusters surpassing a threshold of alpha = 0.05 [corrected for multiple comparisons] are presented). Identified visual areas (V1, V2, MT+, etc.) are from the Colin atlas database. The borders represent Brodmann areas from the Colin atlas. For abbreviations see Figure 2.

also underlie the importance of congruency for multisensory integration already indicated by behavioral experiments (e.g., Meyer and others 2005).

The fact that the association cortex in the STG, the parietal, and superior frontal cortex but also several visual areas respond more strongly to congruent audiovisual stimuli indicates that these areas are also involved in audiovisual integration. This is partly in concordance with the Lewis and others (2000) who found the IPS and the anterior midline to be involved in

multisensory interaction processes. But supplementary to their results we also found evidence for multisensory enhancement in the SMG, SPL, STG, and early visual areas.

The activation in the visual areas could be due either to multisensory interaction at an early visual level or to feedback from higher areas like the parietal association cortex. Classically, multisensory integration is supposed to occur on a relative late stage of the sensory hierarchy (Felleman and Van Essen 1991). But several recent results suggest that also early sensory

cortices might be involved in multisensory interaction and integration processes (Giard and Peronnet 1999; Foxe and others 2000; Macaluso and others 2000; Shams and others 2001; Bhattacharya and others 2002; Falchier and others 2002; Molholm and others 2002; Murray and others 2004).

The area middle temporal (MT)+ was not significantly activated in the condition with coherent visual motion. One possible explanation might be that the difference in stimulus coherence of 16% was not enough to elicit significant differences in the BOLD response. Rees and others (2000) found that the population responses in human V5 increases linearly with stimulus coherence. According to this a rather small signal change would be expected that may be further diminished by averaging over subjects.

Possible Effects of Attentional Processes and Suppression of Eye Movements

For the conditions with 16% coherent visual motion in both conditions with moving acoustic noise (in phase and antiphase), but not for the condition with stationary acoustic noise, several cortical regions, particularly the SPL, SMG, SFG but also extrastriate visual areas, were more strongly activated. Because the conditions with the moving sound source were also more demanding (in which subjects were required to evaluate the relative direction of the visual coherent motion), it is likely that these conditions required more attention to the visual stimuli. Therefore, the enhanced activation in the conditions with 16% coherent visual motion could be at least partly caused by attentional processes. There exists already substantial experimental evidence that attention to visual stimuli or visual motion enhances the neural responses in extrastriate areas and in the intraparietal sulcus (Spitzer and others 1988; Beauchamp and others 1997; O'Craven and others 1997; Buechel and others 1998; Chawla and others 1999).

The pronounced activity in the in phase condition compared with the antiphase condition cannot be explained by enhanced attention to visual motion in general because both in phase and antiphase conditions required subjects to evaluate the direction of the visual coherent motion. However, the stronger activity in the in phase condition could be, besides being caused by multisensory integration processes, explained by enhanced visual processing due to the fact that in the in phase condition the auditory stimuli may direct attention to the direction of the coherent dot movement. It has been shown (Sekuler and Ball 1977) that attention to a specific direction can enhance detection of visual stimuli moving in the attended direction. Furthermore, attention to a specific direction leads also to a more pronounced motion-after effect (Chaudhuri 1990; von Grunau and others 1998; Alais and Blake 1999). In recent event-related potential (ERP) studies (Beer and Roder 2004, 2005), attention directed to a particular direction of motion enhanced processing of both visual and auditory stimuli. However, it is not clear if subjects in our study were using the sound source as a direction cue or if the enhanced BOLD response is due to multisensory integration. Some subjects reported that the motion detection task was easier in the in phase condition, but none of them reported that they were using the auditory motion stimuli deliberately as a cue. This strategy seems reasonable because the subjects knew that the auditory stimuli had no predictive value regarding the direction of the visual motion.

For the conditions with 16% coherent visual motion (in phase, antiphase, and stationary) we found significantly more activity in the dorsal precentral gyrus and the SFG, which may correspond to the frontal eye fields. It has already been shown that when humans attentively pursue objects without moving their eyes, the frontal eye fields are involved, which is also true for the MT+ complex and the area around the intraparietal sulcus (Culham and others 1998). However, the frontal eye field is activated in tasks demanding spatial working memory, as well (Jonides and others 1993; McCarthy and others 1994; Courtney and others 1996; LaBar and others 1999). In the conditions with 16% coherent visual motion, subjects had to judge both the visual and the auditory motion and in the conditions without coherent visual motion they only had to judge the visual component of the stimuli. It is therefore likely that, when subjects had to compare the visual and auditory motion, it was a more taxing task, demanding more cortical activity. The brain activity in the frontal eye fields can be at least partly due to the suppression of eye movements when directional moving stimuli are present (Sheliga and others 1995; Law and others 1997; Petit and others 1999). Because we monitored eye movements during fMRI we could rule out a possible role of saccadic or pursuit intrusions.

In the ventral precentral gyrus we found spatially separated activations in all 3 auditory conditions. This region likely corresponds to the premotor cortex, which is, at least for primates, known to be somatotopically organized (Godschalk and others 1995; Raos and others 2003). Because subjects had to respond with a different finger in the 3 conditions, the differential activations are likely due to response preparation (Simon and others 2002).

However, despite possible functional overlap with attentional processes and fixation control, the results of our study gives effectual evidence for a participation of lateral parietal, superior temporal, and superior frontal cortex in the integration of moving audiovisual stimuli. Because congruent audiovisual motion leads to far more extensive activation patterns, this supports the idea that auditory and visual motion are processed conjointly in these brain areas. Our findings, though promising, have limitations inherent to the utilization of a sparse imaging design. With this approach we are unable to investigate the temporal time course of cortical activation related to specific components of the tasks.

Conclusion

We found that the superior temporal cortex, the SMG, and the superior parietal lobule underlie our ability to integrate audiovisual motion cues and that these 2 regions exhibit a differential sensitivity to in phase and antiphase combinations of audiovisual motion stimuli. Areas in the frontal cortex appear to mediate the integrative and attentive aspects of the task.

Notes

Address correspondence to Mark W. Greenlee, Institute for Experimental Psychology, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany. Email: mark.greenlee@psychologie.uni-regensburg.de.

References

Ahissar M, Ahissar E, Bergman H, Vaadia E. 1992. Encoding of sound-source location and movement: activity of single neurons and interactions between adjacent neurons in the monkey auditory cortex. J Neurophysiol 67:203-215.

Alais D, Blake R. 1999. Neural strength of visual attention gauged by motion adaptation. Nat Neurosci 2:1015-1018.

Andersen RA. 1997. Multimodal representation of space in the posterior parietal cortex and its use in planning movements. Annu Rev Neurosci 20:303-330.

Baumgart F, Gaschler-Markefski B, Woldorff MG, Heinze H, Scheich H. 1999. A movement-sensitive area in auditory cortex. Nature 400:724-726.

Beauchamp MS, Cox RW, DeYoe EA. 1997. Graded effects of spatial and featural attention on human area MT and associated motion processing areas. J Neurophysiol 78:516-520.

Beer AL, Roder B. 2004. Attention to motion enhances processing of both visual and auditory stimuli: an event-related potential study. Cogn Brain Res 18:205-225.

Beer AL, Roder B. 2005. Attending to visual or auditory motion affects perception within and across modalities: an event-related potential study. Eur J Neurosci 21:1116-1130.

Belin P, Zatorre RJ, Hoge R, Evans AC, Pike B. 1999. Event-related fMRI of the auditory cortex. Neuroimage 10:417-429.

Benevento LA, Fallon J, Davis BJ, 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.

Bhattacharya J, Shams L, Shimojo S. 2002. Sound-induced illusory flash perception: role of gamma band responses. Neuroreport 13:1727-1730.

Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Rao SM, Cox RW. 1999. Conceptual processing during conscious resting state: a functional MRI study. J Cogn Neurosci 11:80-95.

Bremmer F, Schlack A, Shah NJ, Zafiris O, Kubischik M, Hoffmann K, Zilles K, Fink GR. 2001. Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. Neuron 29:287-296.

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

Buechel C, Josephs O, Rees G, Turner R, Frith CD, Friston KJ. 1998. The functional anatomy of attention to visual motion: a functional MRI study. Brain 121:1281-1294.

Bushara KO, Weeks RA, Ishii K, Catalan M-J, Tian B, Rauschecker JP, Hallet M. 1999. Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. Nat Neurosci 2:759-766.

Calvert GA, Campbell R, Brammer MJ. 2000. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr Biol 10:649-657.

Calvert GA, Spence C, Stein BE. 2004. The handbook of multisensory processes. Cambridge, England: Bradford Book.

Chaudhuri A. 1990. Modulation of the motion aftereffect by selective attention. Nature 344:60-62.

Chawla D, Rees G, Friston KJ. 1999. The physiological basis of attentional modulation in extrastriate visual areas. Nature 7:671-676.

Colby CL, Duhamel J, Goldberg ME. 1996. Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. J Neurophysiol 76:2841-2852.

Courtney SM, Ungerleider LG, Keil K, Haxby JV. 1996. Object and spatial visual working memory activate separate neural systems in human cortex. Cereb Cortex 6:39-49.

Culham JC, Brandt SA, Cavanagh P, Kanwisher NG, Dale AM, Tootell RB. 1998. Cortical fMRI activation produced by attentive tracking of moving targets. J Neurophysiol 80:2657-2670.

Falchier A, Clavagnier S, Barone P, Kennedy H. 2002. Anatomical evidence of multimodal integration in primate striate cortex. J Neurosci 22:5749-5759.

Felleman DJ, Van Essen DC. 1991. Distributed hierarchical processing in primate visual cortex. Cerebral Cortex 1:1-47.

Foxe JJ, Morocz IA, Murray MM, Higgins BA, Javitt DC, Schroeder CE. 2000. Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. Cogn Brain Res 10:77-83.

Friston KJ, Ashburner J, Poline JB, Frith CD, Heather JD, Frackowiak RSJ. 1995a. Spatial registration and normalization of images. Hum Brain Mapp 2:165-189.

Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ. 1995b. Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp 2:189-210.

Fujisaki W, Nishida S. 2005. Temporal frequency characteristics of synchrony-asynchrony discrimination of audio-visual signals. Exp Brain Res 166:455-464.

Giard MH, Peronnet F. 1999. Auditory-visual integration during multi-modal object recognition in humans: a behavioral and electrophysiological study. J Cogn Neurosci 11:473-490.

Godschalk M, Mitz AR, van Duin B, van der Burg H. 1995. Somatotopy of monkey premotor cortex examined with microstimulation. Neurosci Res 23:269-279.

Gondan M, Roder B. 2006. A new method for detecting interactions between the senses in event-related potentials. Brain Res 1073-1074:389-397.

Griffiths TD, Rees G, Rees A, Green GGR, Witton C, Rowe D, Büchel C, Turner R, Frackowiak RSJ. 1998. Right parietal cortex is involved in the perception of sound movement in humans. Nat Neurosci 1:74-79.

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW. 1999. "Sparse" temporal sampling in auditory fMRI. Hum Brain Mapp 7:213-223.

Haxby JV, Horwitz B, Ungerleider LG, Maisog JM, Pietrini P, Grady CL. 1994. The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. J Neurosci 14:6336-6353.

Holmes AP. 1994. Statistical issues in functional brain mapping [Doctoral dissertation]. University of Glasgow.

Holmes AP, Blair RC, Watson JDG, Ford I. 1996. Non-parametric analysis of statistic images from functional mapping experiments. J Cereb Blood Flow Metab 16:7-22.

Hyvarinen J, Shelepin Y. 1979. Distribution of visual and somatic functions in the parietal associative area 7 of the monkey. Brain Res 169:561-564.

Jonides J, Smith EE, Koeppe RA, Awh E, Minoshima S, Mintun MA. 1993. Spatial working memory in humans as revealed by PET. Nature 363:623-625.

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

Kimmig H, Greenlee MW, Huethe F, Mergner T. 1999. MR-Eyetracker: a new method for eye movement recording in functional resonance imaging. Exp Brain Res 126:443-449.

LaBar KS, Gitelman DR, Parrish TB, Mesulam M. 1999. Neuroanatomic overlap of working memory and spatial attention networks: a functional MRI comparison within subjects. Neuroimage 10:695-704.

Lancaster JL, Summerln JL, Rainey L, Freitas CS, Fox PT. 1997. The Talairach Daemon, a database server for Talairach atlas labels. Neuroimage 5:633.

Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT. 2000. Automated Talairach atlas labels for functional brain mapping. Hum Brain Mapp 10:120-131.

Law I, Svarer C, Holm S, Paulsen OB. 1997. The activation pattern in normal humans during suppression, imagination and performance of saccadic eye movements. Acta Physiol Scand 161:419-434.

Lewis JW, Beauchamp MS, DeYoe EA. 2000. A comparison of visual and auditory motion processing in human cerebral cortex. Cereb Cortex 10:873-888.

Lewis JW, Van Essen DC. 2000. Cortico-cortical connections of visual, sensorimotor, and multimodal processing. Neuroimage 7:378.

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

Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19:1233–1239.

McCarthy G, Blamire AM, Puce A, Nobre AC, Bloch G, Hyder F, Goldman-Rakic P, Shulman RG. 1994. Functional magnetic resonance imaging of human prefrontal cortex activation during a spatial working memory task. Proc Natl Acad Sci USA 91:8690–8694.

Meyer GF, Wuerger SM, Röhrbein F, Zetzsche C. 2005. Low-level integration of auditory and visual motion signals requires spatial co-localisation. Exp Brain Res 166:538–547.

Miller J. 1982. Divided attention: evidence for coactivation with redundant signals. Cognit Psychol 14:247–279.

Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, Foxe JJ. 2002. Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. Cogn Brain Res 14:115–128.

Morein-Zamir S, Soto-Faraco S, Kingstone A. 2003. Auditory capture of vision: examining temporal ventriloquism. Cogn Brain Res 17:154–163.

Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, Schroeder CE, Foxe JJ. 2004. Grabbing your ear: rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. Cereb Cortex 15:963–974.

O'Craven KM, Rosen BR, Kwong KK, Treisman A, Savoy RL. 1997. Voluntary attention modulates fMRI activity in human MT-MST. Neuron 18:591–598.

Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD. 2004. The processing of temporal pitch and melody information in auditory cortex. Neuron 36:767–776.

Pavani F, Macaluso E, Warren JD, Driver J, Griffiths TD. 2002. A common cortical substrate activated by horizontal and vertical sound movement in human brain. Curr Biol 12:1584–1590.

Petit L, Dubois S, Tzourio N, Dejardin S, Crivello F, Michel C, Etard O, Denise P, Roucoux A, Mazoyer B. 1999. PET study of human foveal fixation system. Hum Brain Mapp 8:28–43.

Raos V, Franchi G, Gallese V, Fogassi L. 2003. Somatotopic organization of the lateral part of area F2 (dorsal premotor cortex) of the macaque monkey. J Neurophysiol 89:1503–1518.

Rees G, Friston K, Koch C. 2000. A direct quantitative relationship between the functional properties of human and macaque V5. Nat Neurosci 3:716–723.

Reulen JPH, Marcus JT, Koops D, de Vries FR, Tiesinga G, Boshuizen K, Bos JE. 1988. Precise recording of eye movement: the IRIS technique Part 1. Med Biol Eng Comput 26:20–26.

Rizzolatti G, Fogassi L, Gallese V. 1997. Parietal cortex: from sight to action. Curr Opin Neurobiol 7:562–567.

Sekuler R, Ball K. 1977. Mental sets alters visibility of moving targets. Science 198:60–62.

Shams L, Kamitani Y, Thompson S, Shimojo S. 2001. Sound alters visual evoked potentials in humans. Neuroreport 12:3849–3852.

Sheliga BM, Riggio L, Rizzolatti G. 1995. Spatial attention and eye movements. Exp Brain Res 105:261–275.

Shulman GL, Corbetta M, Buckner RL, Raichle ME, Fiez JA, Miezin FM, Petersen SE. 1997. Top-down modulation of early sensory cortex. Cereb Cortex 7:193–206.

Simon SR, Meunier M, Piettre L, Berardi AM, Segebarth CM, Boussaoud D. 2002. Spatial attention and memory versus motor preparation: premotor cortex involvement as revealed by fMRI. J Neurophysiol 88:2047–2057.

Soto-Faraco S, Lyons J, Gazzaniga M, Spence C, Kingstone A. 2002. The ventriloquist in motion: illusory capture of dynamic information across sensory modalities. Cogn Brain Res 14:139–146.

Soto-Faraco S, Spence C, Kingstone A. 2004. Cross-modal dynamic capture: congruence effects in the perception of motion across sensory modalities. J Exp Psychol Hum Percept Perform 30: 330–345.

Spitzer H, Desimone R, Moran J. 1988. Increased attention enhances both behavioral and neuronal performance. Science 240:338–340.

Stein BE, Meredith MA. 1993. The merging of the senses. Cambridge, MA: MIT Press.

Van Essen DC. 2002. Windows on the brain: the emerging role of atlases and databases in neuroscience. Curr Opin Neurobiol 12:574–579.

Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH. 2001. An integrated software suite for surface-based analyses of cerebral cortex. J Am Med Inform Assoc 8:443–459.

von Grunau MW, Bertone A, Pakneshan P. 1998. Attentional selection of motion states. Spat Vis 11:329–347.

Wagner H, Kautz D, Poganiatz I. 1997. Principles of acoustic motion detection in animals and man. Trends Neurosci 20:583–588.

Wickens TD. 2001. Elementary signal detection theory. Oxford, England: Oxford University Press.