Behavioral/Systems/Cognitive

Audiovisual Synchrony Improves Motion Discrimination via Enhanced Connectivity between Early Visual and Auditory Areas

Richard Lewis and Uta Noppeney

Max Planck Institute for Biological Cybernetics, 72076 Tuebingen, Germany

Audiovisual synchrony enables integration of dynamic visual and auditory signals into a more robust and reliable multisensory percept. In this fMRI study, we investigated the neural mechanisms by which audiovisual synchrony facilitates shape and motion discrimination under degraded visual conditions. Subjects were presented with visual patterns that were rotated by discrete increments at irregular and unpredictable intervals while partially obscured by a dynamic noise mask. On synchronous trials, each rotation coincided with an auditory click. On asynchronous trials, clicks were noncoincident with the rotational movements (but with identical temporal statistics). Subjects discriminated shape or rotational motion profile of the partially hidden visual stimuli. Regardless of task context, synchronous signals increased activations bilaterally in (1) calcarine sulcus (CaS) extending into ventral occipitotemporal cortex and (2) Heschl's gyrus extending into planum temporale (HG/PT) compared with asynchronous signals. Adjacent to these automatic synchrony effects, synchrony-induced activations in lateral occipital (LO) regions were amplified bilaterally during shape discrimination and in the right posterior superior temporal sulcus (pSTS) during motion discrimination. Subjects' synchrony-induced benefits in motion discrimination significantly predicted blood oxygenation level-dependent synchrony effects in V5/hMT+. According to dynamic causal modeling, audiovisual synchrony increased connectivity between CaS and HG/PT bidirectionally, whereas shape and motion tasks increased forwards connectivity from CaS to LO or to pSTS, respectively. To increase the salience of partially obscured moving objects, audiovisual synchrony may amplify visual activations by increasing the connectivity between low level visual and auditory areas. These automatic synchrony-induced response amplifications may then be gated to higher order areas according to behavioral relevance and task context.

Introduction

A creature moving through the undergrowth and partially obscured by intervening foliage may be more readily identified when each footfall is accompanied by a coincident sound, heralding motion of target between occlusions and background visual noise. Temporal proximity between auditory and visual stimuli is an essential cue, informing the brain whether they emanate from a common source and should hence be integrated in perception (Munhall et al., 1996; Sekuler et al., 1997; Shams et al., 2002). Synchrony drives integration of signals from multiple senses, enabling behavioral benefits, as demonstrated by psychophysical studies showing enhanced intelligibility of synchronous audiovisual speech (Pandey et al., 1986) and improved detection of simple visual targets when accompanied by synchronous tones (Vroomen and de Gelder, 2000; Frassinetti et al., 2002; Noesselt et al., 2008; Olivers and Van der Burg, 2008).

Since the seminal neurophysiological studies showing temporal concordance as a determinant of multisensory integration

(Stein and Meredith, 1993), several human imaging studies have identified audiovisual synchrony effects for nonspeech stimuli in a widely distributed neural system encompassing the superior colliculus, insula, occipital, inferior parietal, and superior/middle temporal regions. However, these previous studies investigated the effects of synchrony on suprathreshold stimulus processing and primarily during passive stimulus exposure (Calvert et al., 2001; Dhamala et al., 2007, Noesselt et al., 2007) or explicit synchrony judgments that confound stimulus and task dimensions (Bushara et al., 2001). Yet the integration of synchronous multisensory cues conveys the greatest behavioral advantage when visual and/or auditory inputs are unreliable (Heron et al., 2004; Ross et al., 2007). Furthermore, the behavioral context in which visual objects in our environment are scrutinized critically influences how they are processed in the brain (Corbetta et al., 1991; Peuskens et al., 2004).

In this fMRI study, we investigated the neural mechanisms by which audiovisual synchrony facilitates shape and motion discrimination in visually obscured scenes. Subjects viewed partially occluded, discretely rotating, target dot patterns accompanied by synchronous or asynchronous acoustic cues. In a visual selective-attention task, they discriminated the shape or rotational motion profile of visual stimuli. This paradigm enabled dissociation of three distinct synchrony-induced effects on brain activations. First, we identified automatic audiovisual synchrony effects emerging regardless of task context. Second, we investigated
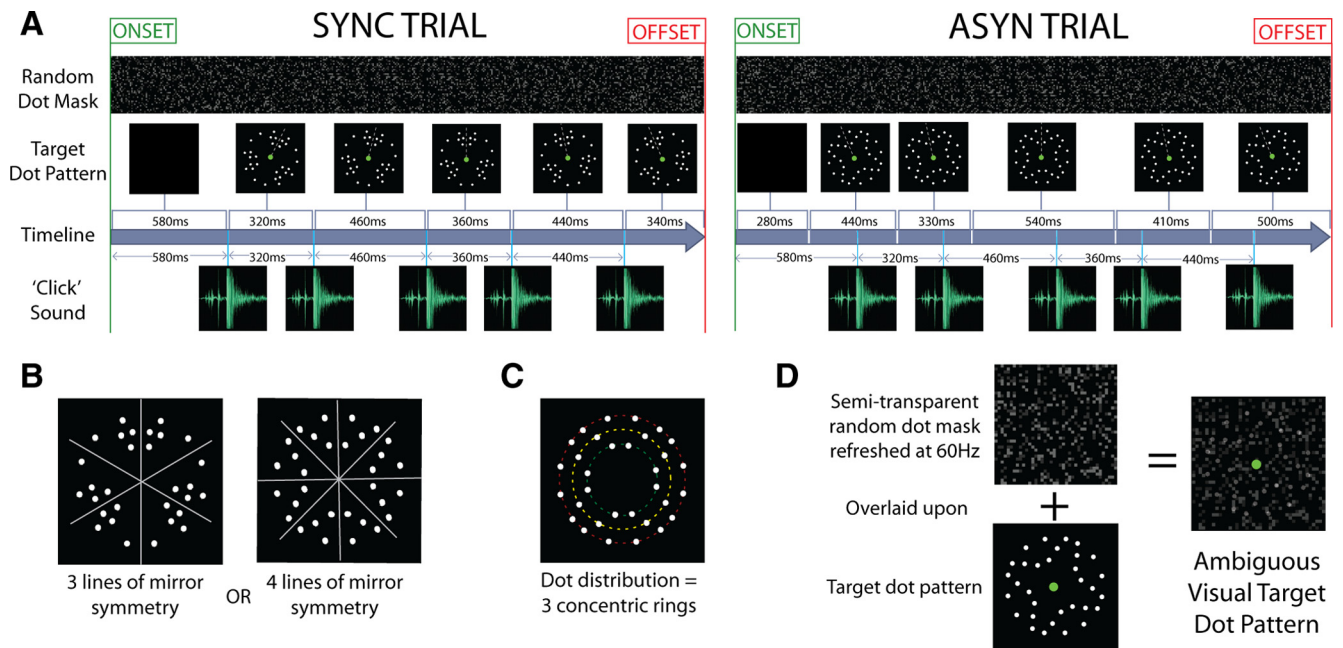
A

Figure 1. Audiovisual stimuli. ***A***, Example trials. Each trial consisted of five consecutively presented, circular dot patterns, with each pattern in-plane rotated relative to the previous one and five acoustic clicks that were emitted synchronously (SYNC trial) or asynchronously (ASYN trial) with respect to the rotation of the visual dot patterns. The target circular dot array was overlaid with a rapid (60 Hz) sequence of partially transparent random dot masks in a dynamic fashion. NB gray dashed lines are added in the figure for illustration purposes only to indicate the unidirectional (e.g., see SYNC trial) or reversing rotation (e.g., see ASYN trial) of sequential circular dot patterns. ***B***, ***C***, Spatial arrangement of visual dot patterns. Visual dot patterns were structured according to three or four lines of mirror symmetry (***B***). The dots were distributed around three concentric rings of diameter 2.02° (green dashes), 2.58° (yellow dashes), and 3.14° (red dashes) visual angle (***C***). ***D***, The target dot patterns were overlaid with partially transparent random noise masks to introduce visual ambiguity.

where synchrony effects were amplified according to behavioral relevance (i.e., synchrony × task interaction). Third, we investigated whether each subject's synchrony-induced behavioral benefit predicted their synchrony-induced blood oxygenation level-dependent (BOLD) effects. Finally, we used dynamic causal modeling to investigate how synchrony effects emerge from distinct interactions among brain regions.

At the behavioral level, we hypothesized that shape and motion discrimination would be improved by audiovisual synchrony, despite auditory stimuli alone conveying no information about visual shape or motion profile. Given accumulating evidence for early audiovisual interactions in sensory-specific cortices (Schroeder and Foxe, 2002, 2005; Foxe and Schroeder, 2005; Ghazanfar and Schroeder, 2006; Kayser and Logothetis, 2007; Driver and Noesselt, 2008), we expected synchrony effects in low-level visual/auditory cortices and further amplification in higher level visual cortices depending on task context. Specifically, we expected synchrony effects to be amplified in area V5/hMT+ for motion discrimination and in ventral occipitotemporal regions for shape discrimination. Furthermore, synchrony effects should be related to subjects' synchrony-induced benefits in shape and motion discrimination.

Materials and Methods

Subjects. Sixteen healthy volunteers (age range, 18–31 years; 2 left-handed; 5 female) participated in this study after giving informed consent. Data from three additional subjects were excluded because either performances in both tasks was >2 SD from the group mean (one subject) or performance difference between shape and motion discrimination was >2 SD from the group mean difference (two subjects). All subjects had normal or corrected-to-normal vision and reported normal hearing. The study was approved by the joint human research review committee of the Max Planck Society and the University of Tuebingen.

Experimental design. Subjects were presented on each trial with a rotation sequence of circular dot patterns that were accompanied by acoustic clicks (Fig. 1A). A sequence consisted of five circular dot patterns with each pattern in-plane rotated relative to the previous one and five acoustic clicks that were emitted synchronously (SYNC trials) or asynchronously (ASYN trials) with respect to the rotation of the visual dot patterns. In a visual selective-attention paradigm, subjects discriminated the spatial arrangement (shape) of the circular dot array or its rotation sequence (motion). Thus, the experimental paradigm conformed to a 2 × 2 factorial design manipulating, as follows: (1) audiovisual synchrony (two levels), wherein the auditory clicks were presented synchronously or asynchronously relative to the onset of each dot pattern in the rotation sequence; and (2) visual task (two levels), whereby during the shape-task subjects discriminated the spatial arrangement of circular dot patterns (Fig. 1B) as triple-symmetric (three lines of mirror symmetry) or quadruple-symmetric (four lines of mirror symmetry), while during the motion task, subjects discriminated the rotation sequence as unidirectional, e.g., only counterclockwise in SYNC trial example (Fig. 1A, left), or reversing, e.g., clockwise followed by counterclockwise in ASYN trial example (Fig. 1A, right).

Identical audiovisual stimuli were presented during shape and motion task conditions. Similarly, identical auditory and visual stimulus components were presented during synchronous and asynchronous conditions with only the temporal relationship (i.e., synchronous vs asynchronous) being different.

In the following, the visual stimuli, the auditory stimuli, and their temporal relationship will be described in more detail.

Visual stimuli. In each trial, the visual stimulus consisted of a sequence of five dot patterns that were each presented for 240–620 ms (see Timing of auditory and visual stimuli for details of stimulus timings). Each random dot pattern was in-plane rotated relative to the previous one by a fixed increment of between 4 and 14 degrees in either a clockwise or counterclockwise direction around the centrally positioned fixation circle. The rotation sequence of a trial was either unidirectional (i.e., rotation within a trial was consistently clockwise or counterclockwise) or
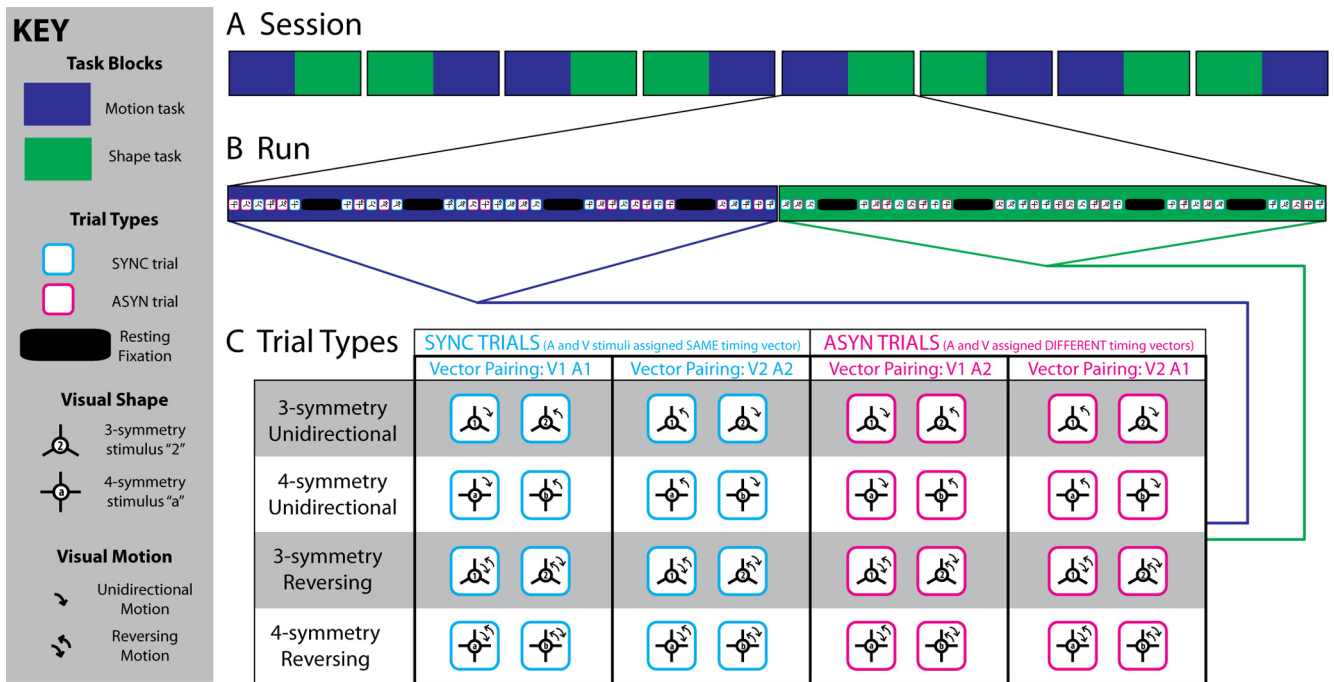
Figure 2. Counterbalancing of task and stimulus trials. ***A***, For eight subjects, each scanning session comprised four runs of motion then shape discrimination alternating with four runs of shape then motion discrimination, and vice versa for the other eight. ***B***, Stimuli encountered during the motion discrimination part of each run were always identical to those of the shape discrimination part of each run (i.e., all stimuli illustrated in ***C*** once under both task conditions), but were presented in a different pseudo-randomized order. ***C***, Synchronous trials occurred when visual and auditory stimuli were driven by the same interstimulus interval vector (i.e., V1A1 or V2A2) and, when visual and auditory stimuli were driven by different interstimulus interval vectors (i.e., V1A2 or V2A1), the trials were asynchronous, thus the timings of audiovisual stimuli were fully counterbalanced across each run. During every run, two of the 16 visual stimuli with three planes of mirror symmetry ("1" and "2") and two of the 16 with four planes of mirror symmetry ("a" and "b") were randomly selected and manipulated so that across all trials each satisfied all possible combinations of the two different motion profiles × two different SYNC temporal profiles × two different ASYN temporal profiles × two different task constraints.

reversing (i.e., two clockwise rotations were followed by two counter-clockwise rotations, or vice versa).

Each of the five frames within a rotation sequence comprised an array of dots arranged in one of 32 different symmetrical configurations. The luminance of the background was 1.04 cd/m^2 and each dot in the array had a luminance of 2.84 cd/m^2. All configurations produced a circular pattern of dots subtending 3.14° in diameter, but excluding the central region of 2° diameter (102 cm viewing distance). Each dot subtended 0.2° visual angle. They were organized in three concentric rings in parafoveal space with diameters 2.02, 2.58, 3.14°, respectively (Fig. 1*C*), thus beyond the field of view of the foveola (∼1.0°) and rod-free fovea (∼1.7°). Dots were distributed around these three rings to create 16 different circular patterns with three lines of mirror symmetry and a further 16 different stimuli with four lines of mirror symmetry (Fig. 1*B*). To render the two visual discrimination tasks more difficult and adjust subjects' performance levels, the visual stimulus was overlaid with a rapid (60 Hz) sequence of random dot noise masks (i.e., 60 different noise-mask frames per second) that partially occluded the target circular dot array (Fig. 1*D*) in a dynamic fashion. Each frame consisted of 0.2 × 0.2° elements in a 6 × 6° grid, with a specific proportion of which (10–36%) filled with opaque gray-scale pixels and those remaining rendered transparent. Half of the opaque-mask dots had the same luminance as the background (1.04 cd/m^2) and the other half were brighter than both background and target dots (5.73 cd/m^2).

Auditory stimuli. In each trial, a sequence of five identical click sounds [50 ms duration, including 5 ms rise/fall; interclick intervals of 240–620 ms; 75 dB sound pressure level (SPL)] were presented synchronously or asynchronously relative to the onsets of the five visual dot patterns.

Timing of auditory and visual stimuli. For SYNC trials, the onsets of the visual circular dot patterns and auditory clicks were determined by a common onset vector (Fig. 1*A*, left) generated by sampling five intervals from a Poisson distribution (lambda = 18, scaling factor = 20; producing individual interstimulus interval values in milliseconds) with the intervals being resampled if they fell outside of the desired range of 240–

620 ms. For ASYN trials, visual and auditory onsets were determined by a pair of independently sampled interval vectors (Fig. 1*A*, right). To prevent subjects from integrating auditory and visual signals that co-occur close in time in the asynchronous case, pairs of auditory and visual onset vectors were eliminated and resampled if they included onsets of visual and auditory events that co-occurred within 120 ms of one another. Pairs of auditory and visual onset vectors were also eliminated and resampled if both their cumulative sums did not fall between 1800 and 2250 ms to ensure that events occurred throughout the majority of the 2.5 s trial duration. This procedure rendered the signal onsets relatively unpredictable, i.e., subjects were unable to anticipate exactly when each visual and auditory stimulus transition might occur during a trial. To ensure that the resampling procedure did not inadvertently introduce predictable interstimulus intervals, we performed Spearman's rank correlation analyses to measure the degree of correlation between successive intervals in each subject. We found that, in every subject, all four correlation analyses (between first and second, second and third, third and fourth, and fourth and fifth intervals) resulted in rho values of between −0.4 and 0.4, i.e., no strong correlation between successive stimulus transition intervals.

To ensure identical timings across conditions, each onset vector from the asynchronous pairings (e.g., onset vector "1" and onset vector "2") was used equally often in any condition (Fig. 2*C*), determining the onsets of both auditory (A) and visual (V) events in SYNC trials (e.g., V1A1, V2A2) and auditory events in one ASYN trial and visual events in another (e.g., V1A2, V2A1). As a result of this counterbalancing procedure, there was no difference in temporal statistics between conditions within each scanning run, apart from the temporal relationship (i.e., synchronous or asynchronous) between auditory and visual streams within a trial.

Stimulus presentation. Visual and auditory stimuli were presented using Cogent (John Romaya, Vision Laboratory, UCL; http://www.vislab.ucl.ac.uk/), running on Matlab 6.5.1 (Mathworks) on an IBM-compatible PC with Windows XP (Microsoft). Visual stimuli were back-projected onto a Plexiglas screen using an LCD projector (JVC) visible to

the subject through a mirror mounted on the MR head coil. Auditory stimuli were presented at ~75 dB SPL using MR-compatible pneumatic headphones (Siemens). Subjects made responses using an MR-compatible custom-built button device connected to the stimulus computer.

Experimental procedure. Each subject participated in one scanning session (Fig. 2*A*) consisting of eight experimental runs (~7 min each interleaved with ~3 min rest). Each run comprised of 16 trials for each condition in our 2 × 2 factorial design (i.e., motion SYNC, motion ASYN, shape SYNC, shape ASYN), amounting to a total of 128 (16 × 8) trials per condition in the entire scanning experiment. The task factor was manipulated in long periods, with each task covering one continuous half of a run (i.e., 32 motion-discrimination trials followed by 32 shape-discrimination trials, or vice versa). The synchronous and asynchronous trials were presented in a pseudo-randomized fashion within each task period (Fig. 2*B*). The activation trials (trial duration, 2.5 s; intertrial interval, 2 s) were presented in blocks of eight trials interleaved with 8 s fixation. A pseudo-randomized trial and stimulus sequence was generated for each subject. The order of task conditions was counterbalanced within and across subjects.

In each run, 4 of the 32 possible concentric dot patterns were randomly selected as visual stimuli (2 of the 16 possible triple-symmetric patterns and 2 of the 16 possible quadruple-symmetric patterns). Each of these four concentric dot patterns was assigned its own unique pair of desynchronized onset vectors ("1" and "2") and were each presented 16 [2 × 2 × 2 × 2 times; i.e., two different SYNC temporal profiles (V1A1 and V2A2), two different ASYN temporal profiles (V1A2 and V2A1), two different task contexts (motion and shape discrimination), and two different motion profiles (unidirectional and reversing)] (Fig. 2*C*). Hence, for the SYNC and ASYN conditions, auditory and visual signals differed only with respect to their temporal relationship, i.e., (a)synchrony, but were otherwise identical. Further, identical stimuli and audiovisual timings were used in both task contexts. In both tasks, subjects indicated their response as accurately as possible via a two-choice key press at the end of each trial. Task instructions were given at the beginning and in the middle of each scanning run via visual display (i.e., a continuous half of each scanning run was dedicated to a single task). The fixation circle (0.22° diameter) also color-coded the current task (blue, motion; green, shape).

Before scanning, subject's individual performance was adjusted outside the scanner to 75–80% accuracy in both the shape and motion tasks by manipulating the proportion (10–36%) of opaque dots in the random-dot noise mask and the angle of rotation (4–14°) using the descending method of limits. These stimulus parameters were tested inside the scanner just before scanning to ensure the desired performance level had been attained. The two parameters were dynamically adjusted between runs to maintain subjects' performance level (N.B.: nevertheless identical stimuli were always presented equally often in all conditions).

Analysis of behavioral data. A two-by-two, factorial, repeated-measurement ANOVA was used to investigate the main effects of stimulus synchrony, task, and their interactions on performance accuracy across subjects. As task-specific performance improvements were expected for audiovisual synchrony, these effects were tested in preplanned *t* tests.

MRI acquisition. A 3 Tesla Siemens Tim Trio system was used to acquire both T1 anatomical volume images [repetition time (TR) = 2300 ms; echo time (TE) = 2.98 ms; inversion time (TI) = 1100 ms; flip angle = 9°; field of view (FOV) = 256 × 240 mm; 176 slices; isotropic spatial resolution, 1 mm³] and T2*-weighted axial echoplanar images with BOLD contrast (GE-EPI; Cartesian *k* space sampling; flip angle = 90°; FOV = 192 × 192 mm; TE = 40 ms; TR = 3080 ms; 38 axial slices, acquired sequentially in ascending direction; spatial resolution 3 × 3 × 3 mm³ voxels; interslice gap, 0.4 mm; slice thickness, 2.6 mm). There were eight runs with a total of 131 volume images per run. The first four volumes of each run were discarded to allow for T1-equilibration effects.

fMRI analysis. The fMRI data were analyzed using the statistical parametric mapping software SPM5 (Wellcome Department of Imaging Neuroscience, London; http//www.fil.ion.ucl.ac.uk/spm). Scans from each subject were realigned using the first as a reference, unwarped, spatially normalized into MNI standard space using parameters derived

from structural unified segmentation, resampled to 3 × 3 × 3 mm³ voxels, and spatially smoothed with a Gaussian kernel of 8 mm full-width at half-maximum. The data were high-pass filtered to 1/128 Hz and a first-order autoregressive model was used to model the remaining serial correlations (Friston et al., 2002). The data were modeled voxelwise using a general linear model that included regressors obtained by convolving each event-related unit impulse (for activation trials) or 8 s boxcar (for fixation periods) with a canonical hemodynamic response function and its first temporal derivative. In addition to modeling the four conditions in the 2 × 2 factorial design and the fixation periods, the statistical model included error trials (pooled over SYNC and ASYN trials but separately for each task). Realignment parameters were included as nuisance covariates to account for residual motion artifacts. Condition-specific effects for each subject were estimated according to the general linear model and passed to a second-level analysis as contrasts. This involved creating the contrast images (pooled, i.e., summed over runs) pertaining to the main and simple main effects of synchrony, task, interaction, and all activation trials > fixation. These contrast images were entered into independent, second-level, one-sample *t* tests. In addition, the simple main effect of synchrony during motion discrimination only (the only task in which a behavioral benefit was observed) was entered into a regression analysis that used each subject's synchrony-induced perceptual benefit as a predictor for their synchrony-induced activation increase.

Inferences were made at the second level to allow a random-effects analysis and inference at the population level (Friston et al., 1999).

Search volume constraints. Each effect was tested for in two nested search volumes. The first search volume included the entire brain. Based on our a priori hypothesis that improvements in visual discrimination may involve increased activation of regions of cortex specialized for processing moving visual stimuli, the second search volume included both left and right V5/hMT+, as defined by cytoarchitectonic probability maps (Eickhoff et al., 2005; Malikovic et al., 2007). Unless otherwise stated, activations are reported at *p* < 0.05 corrected at the cluster level for multiple comparisons within the entire brain or the V5/hMT+ volume using an auxiliary (uncorrected) voxel threshold of *p* < 0.001. The auxiliary threshold defines the spatial extent of activated clusters, which form the basis of our (corrected) inference. We report all clusters >50 voxels when correcting for the entire brain and all clusters >1 voxel when correcting for V5/hMT+ volume.

Results of the random-effects analysis are superimposed onto a structural image created by averaging the subjects' normalized images.

Dynamic causal modeling. Dynamic causal modeling (DCM) treats the brain as a dynamic input–state–output system (Friston et al., 2003; Penny et al., 2004a, 2004b). The inputs correspond to conventional stimulus functions encoding experimental manipulations. The state variables are neuronal activities and the outputs are the regional hemodynamic responses measured with fMRI. The idea is to model changes in the states, which cannot be observed directly, using the known inputs and outputs. Critically, changes in the states of one region depend on the states (i.e., activity) of others. This dependency is parameterized by effective connectivity. There are three types of parameters in a DCM, as follows: (1) input parameters, which describe how much brain regions respond to experimental stimuli; (2) intrinsic parameters, which characterize effective connectivity among regions; and (3) modulatory parameters, which characterize changes in effective connectivity caused by experimental manipulation. This third set of parameters allows us to explain, for instance, the increased activation for synchronous trials by changes in coupling among brain areas. Importantly, this coupling (effective connectivity) is expressed at the level of neuronal states. DCM employs a forward model, relating neuronal activity to fMRI data that can be inverted during the model fitting process. Put simply, the forward model is used to predict outputs using the inputs. The parameters are adjusted (using gradient descent) so that the predicted and observed outputs match. This adjustment corresponds to the model-fitting.

For each subject, seven DCMs were constructed. Each DCM included four regions, as follows: (1) the right calcarine sulcus (CaS) as the likely site of the primary visual cortex showing the main effect of synchrony (CaS: +15, −96, −3); (2) the right Heschl's gyrus (HG)/planum tempo-

rale (PT) as an early auditory area showing the main effect of synchrony (HG/PT: +51, −21, +6); (3) the right posterior superior temporal sulcus (pSTS), identified by a motion × synchrony interaction (pSTS: +63, −36, +9); and (4) a right lateral occipital (LO) region exhibiting a shape × synchrony interaction (LO: +42, −69, −12). The four regions were bidirectionally connected. The timings of the onsets were individually adjusted for each region to match the specific time of slice acquisition (Kiebel et al., 2007). Visual stimuli entered as extrinsic inputs to CaS and auditory stimuli to HG/PT. Holding the intrinsic and extrinsic connectivity structure constant, the seven DCMs manipulated the connection(s) that was (were) modulated by synchrony, whereas the connections modulated by task were fixed (see Fig. 6 for candidate DCMs).

The regions were selected using the regional subcluster maxima of the relevant contrasts from our whole brain random-effects analysis. For consistency, we constrained the DCM to regions in the right hemisphere. Region-specific time-series (concatenated over the eight runs and adjusted for confounds) comprised the first eigenvariate of all voxels within a sphere of 4 mm radius centered on the peak subcluster voxel from each anatomical region as identified by the relevant second-level contrast (or directly adjacent to the peak voxel if it was on the cluster border).

Bayesian model comparison. To determine the most likely of the seven DCMs given the observed data from all subjects (i.e., the model with the highest posterior probability), we implemented a fixed-effects group analysis (Penny et al., 2010). The model posterior probability was obtained by combining the model evidences of each subject-specific DCM and the model prior according to Bayes rule (Penny et al., 2004a). The model evidence as approximated by the free energy does not only depend on model fit but also model complexity. The posterior probability of a model family is the sum of the posterior probabilities of all models within a family. The direct family 1 included models A, D, E, and G, in which synchrony was allowed to modulate bilateral connectivity between CaS and HG/PT directly. The indirect family 2 comprised models B, C, and F, which did not allow synchrony to modulate these direct intrinsic connections. Please note that the direct models do not imply that the interactions between HG/PT and CaS are mediated monosynaptically; they may well be mediated polysynaptically through other omitted brain regions. However, higher evidence for the direct relative to the indirect family suggests that there are effective interactions between HG/PT and CaS that cannot be explained by influences from the other areas modeled (Friston, 2009; Roebroeck et al., 2009).

For the optimal model, the subject-specific, modulatory, extrinsic and intrinsic connection strengths were entered into *t* tests at the group level. This allowed us to summarize the consistent findings from the subject-specific DCMs using classical statistics. Model comparison and statistical analysis of connectivity parameters of the optimal model were used to investigate how synchrony may induce activation increases by modulating the coupling among several brain areas.

Results

Behavioral data

For performance accuracy, a two-by-two, factorial, repeated-measurement ANOVA with the factors synchrony (SYNC, ASYN) and task (motion, shape) identified a significant main effect of synchrony ($F_{(1,15)} = 30.1$; $p < 0.001$), yet no main effect of task ($F_{(1,15)} = 0.1$; $p = 0.76$); hence activation differences between shape and motion tasks cannot be attributed to task difficulty confounds. Discrimination performances for each condition (across subject mean) were as follows: SYNC shape task = 78%; ASYN shape task = 78%; SYNC motion task = 81%; ASYN motion = 75% (NB without exclusion of the three subjects, nearly equivalent performance values were obtained i.e., SYNC shape = 79%, ASYN shape = 79%, SYNC motion = 81%, ASYN motion = 76%). There was a significant interaction between synchrony and task ($F_{(1,15)} = 13.3$; $p = 0.002$); follow-up Student's *t* tests on the simple main effects of synchrony showed that subjects showed a significant increase (6.2% across-subject mean) in performance accuracy for synchronous relative to asynchronous
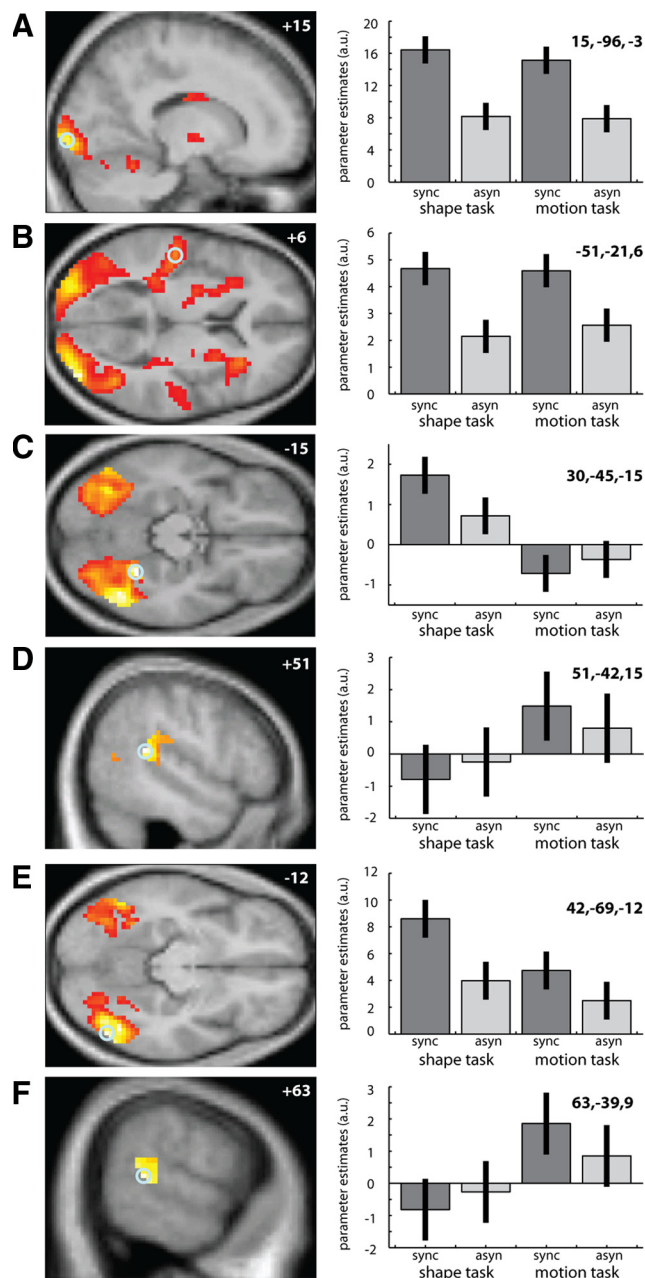
Figure 3. Main effects of synchrony, task and interactions. Left, Main effect of synchrony (*A*, *B*), main effect of task (*C*, *D*), and task-dependent synchrony-enhanced activations (*E*, *F*) are displayed on sagittal and transverse sections of a normalized mean structural image. Right, Parameter estimates (across subjects' mean ± confidence interval) for the regressors indexing the synchronous (dark gray) and asynchronous (light gray) conditions during shape and motion discrimination within the single voxel at the given coordinates (circled voxel). The bar graphs represent the value of each parameter estimate in nondimensional, arbitrary units (a.u.; related to percentage whole brain mean). *A*, *B*, Increased activations for synchronous relative to asynchronous trials ($p < 0.001$, uncorrected; voxels > 50) in primary visual area within the calcarine sulcus (*A*) and primary auditory area within Heschl's gyrus (*B*). *C*, *D*, Increased activations ($p < 0.001$, uncorrected; voxels > 50) for shape task relative to motion task in anterior fusiform gyrus (*C*) and for motion task relative to shape task in posterior superior temporal sulcus (*D*). *E*, *F*, Significant interactions ($p < 0.001$, uncorrected; voxels > 50) between task and synchrony. Synchrony effects are enhanced during shape discrimination in lateral occipital regions (*E*) and during motion discrimination in the posterior superior temporal sulcus (*F*).

trials for motion ($t = 6.93$; $p < 0.001$), but not for shape discrimination ($t = 0.09$; $p = 0.47$). Reaction time data were not statistically analyzed, as subjects were instructed to respond after the full sequence of audiovisual events in each trial had finished.

Table 1. fMRI results

Contrast/regions	Coordinates (MNI space)			Z score	Cluster size (number of voxels)	p Value (cluster level corrected)
Main effects						
Audiovisual synchrony (SYNC)						
Left inferior occipital gyrus	−30	−87	−9	6.50	3517	$p < 0.001$
Right middle occipital gyrus	39	−87	6	6.37		
Right CaS	15	−96	−3	5.97		
Left CaS	−9	−96	−12	5.90		
Left V5/hMT+	−45	−72	−6	5.63	24	$p < 0.001$*
Right V5/hMT+	45	−66	0	5.37	20	$p < 0.001$*
Left precentral gyrus	−36	−27	69	4.72	339	$p < 0.001$
Left superior parietal gyrus	−36	−48	63	4.62		
Left putamen	−24	−6	12	4.69	752	$p < 0.001$
Left HG/PT	−51	−21	6	4.67		
Right putamen	21	−9	12	4.62	375	$p < 0.001$
Right anterior insula	27	21	0	4.61		
Right superior temporal gyrus	54	−33	12	4.19	154	$p < 0.001$
Right HG/PT	51	−21	9	3.69		
Shape task						
Right fusiform gyrus	30	−45	−15	6.40	1977	$p < 0.001$
Right ventral occipitotemporal cortex	48	−60	−12	6.09		
Left IPS	−24	−60	39	5.67	1647	$p < 0.001$
Left ventral occipitotemporal cortex	−36	−69	−12	5.62		
Left precentral sulcus	−45	3	33	4.81	74	$p < 0.001$
Right IPS (horizontal segment)	51	−30	42	4.18	55	$p < 0.05$
Motion task						
Left posterior supramarginal gyrus	−63	−42	27	5.42	458	$p < 0.001$
Left pSTG	−63	−36	18	4.62		
Right pSTS	51	−42	15	5.18	1033	$p < 0.001$
Right pSTG	63	−33	15	4.59		
Left superior frontal gyrus	−12	−3	66	4.36	172	$p < 0.001$
Right precentral gyrus	57	−6	45	4.21	68	$p < 0.05$
Right V5/hMT+	51	−63	9	3.26	1	$p < 0.05$*
Interactions						
SYNC × shape task						
Right ventral occipitotemporal cortex	48	−60	−12	6.24	1079	$p < 0.001$
Right LO region	42	−69	−12	5.86		
Left ventral occipitotemporal cortex	−48	−60	−12	5.34	312	$p < 0.05$
Left fusiform gyrus	−33	−51	−12	4.58		
Left lateral occipital regions	−33	−69	−15	4.50		
Left IPS	−24	−66	48	5.17	524	$p < 0.05$
Left middle occipital gyrus	−30	−93	15	4.63		
Left superior parietal gyrus	−24	−60	57	4.49		
SYNC × motion task						
Right pSTG	51	−39	27	5.21	632	$p < 0.001$
Right pSTS (upper bank)	63	−42	9	4.57		
Left pSTS	−60	−54	6	4.43	164	$p < 0.001$
Right middle frontal gyrus	27	48	15	4.23	88	$p < 0.001$
Regression analysis						
SYNC-enhanced activations significantly predicted by subjects, performance improvements (motion task only)						
Right V5/MT+	54	−66	6	3.25	3	$p < 0.05$*

*Small volume correction at cluster level, corrected for volume of bilateral V5/hMT+, as defined by cytoarchitectonic maps.

CaS, Calcarine sulcus; HG/PT, Heschl's gyrus/planum temporale; pSTS, posterior superior temporal sulcus; IPS, intraparietal sulcus; pSTG, posterior superior temporal gyrus.

fMRI: conventional analysis

This analysis tested for the main effect of synchrony, the main effect of task, and the interaction between synchrony and task. Unless otherwise stated, activations are reported corrected ($p < 0.05$) for multiple comparisons within the entire brain. To focus on the neural systems involved in audiovisual processing, we report in our figures all activations within the neural systems activated for stimulus processing relative to fixation (using the inclusive masking option at $p < 0.2$, uncorrected for purposes of illustration only). MNI coordinates of the central/most active voxel are reported in figures as follows: 51 (x), −21 (y), 6 (z).

Main effect of synchrony

SYNC relative to ASYN trials increased activation bilaterally in both visual areas [extending from the calcarine sulcus (Fig. 3A) into middle/inferior occipital gyrus and V5/hMT+] and in auditory areas [Heschl's gyrus and planum temporale (Fig. 3B)]. Bilateral activations were also observed in the parietal cortex, putamen, and anterior insula (Table 1). Since identical auditory and visual stimuli were presented in synchronous and asynchronous conditions, response amplification cannot be attributed to differences in the auditory or visual signals themselves, but can only be accounted for by differences in the temporal relationship
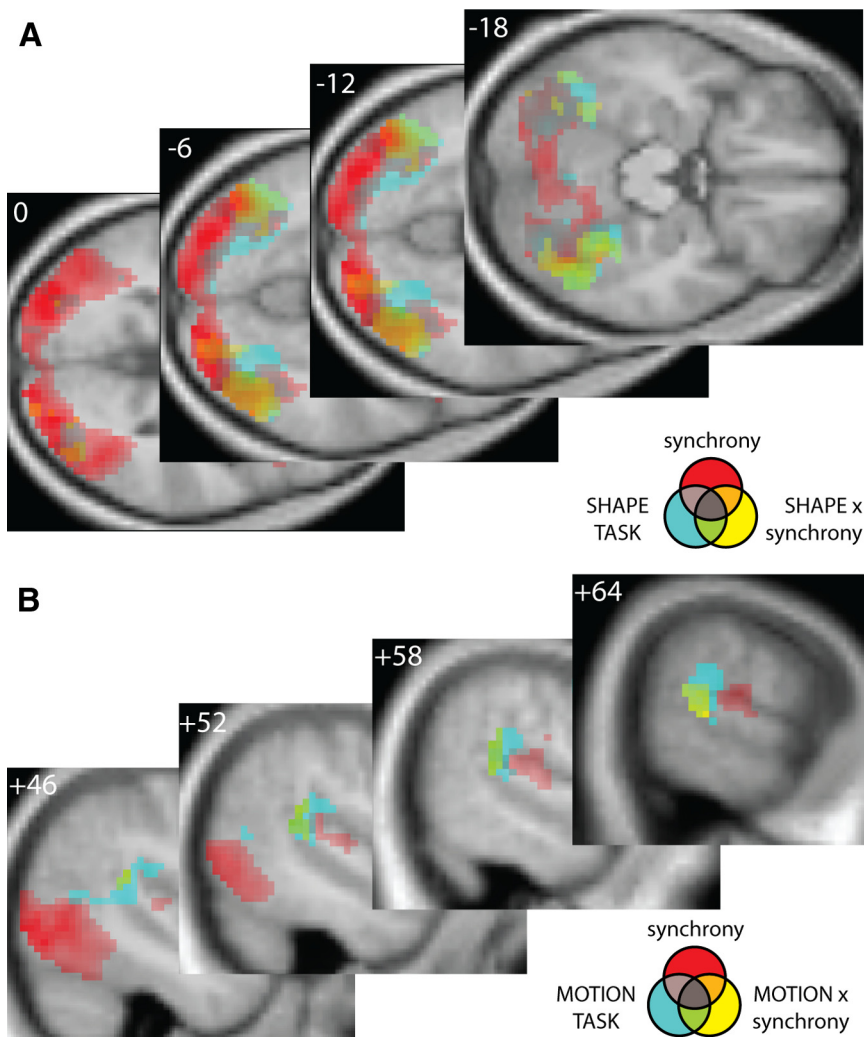
Figure 4. **A**, **B**, Main effects of audiovisual synchrony, task, and interactions between task and synchrony on transverse (**A**) and sagittal (**B**) sections of a normalized mean structural image ($p < 0.001$, uncorrected; voxels > 50). **A**, Red, Main effect of audiovisual synchrony. Cyan, Shape > motion discrimination. Yellow, Synchrony effects that are increased for shape discrimination. Orange, Common to main effect of synchrony and interaction in lateral occipital regions. Green, Common to main effect of shape task and interaction in fusiform and lingual gyri. **B**, Red, Main effect of audiovisual synchrony. Cyan, Motion > shape discrimination. Yellow, Synchrony effects that are increased for motion discrimination. Green, Common to main effect of motion task and interaction contrasts in pSTS.
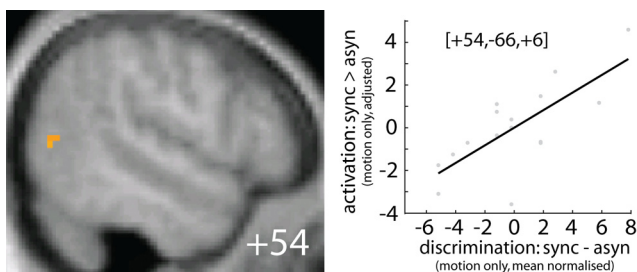
Figure 5. Behavioral relevance of audiovisual synchrony effects. Left, Synchrony effects predicted by subject-specific synchrony-induced improvements in motion discrimination within right V5/MT+ displayed on sagittal sections of a normalized mean structural image. Right, Scatter plot depicts the regression of the regional (adjusted) fMRI signal to subject-specific synchrony-induced performance improvements. The ordinate represents (adjusted) fMRI signal at $x = 54$, $y = -66$, $z = 6$ and the abscissa represents subject's synchrony-induced improvement during motion discrimination (after mean correction).

between the auditory and visual signals, i.e., audiovisual synchrony. As shown in the parameter estimate plots (Fig. 3 *A*,*B*, right), activation increases for synchronous trials were observed regardless of whether subjects actively discriminated visual shape or motion, pointing to a low-level, automatic mechanism of synchrony-induced response amplification. The opposite contrast, i.e., ASYN relative to SYNC trials, revealed no significant activations after correction for multiple comparisons.

Main effect of task

Shape relative to motion discrimination increased activity across bilateral ventral occipitotemporal cortex peaking in the anterior fusiform gyrus (Figs. 3*C*, 4*A*) and extending posterolaterally into regions with coordinates corresponding to those for area LO, known to be involved in object perception (Malach et al., 1995; Grill-Spector et al., 1998; Sarkheil et al., 2008). Activations were also observed bilaterally in the intraparietal and precentral sulci (Table 1).

In contrast, motion relative to shape discrimination enhanced activations in right V5/hMT+, bilaterally in posterior superior temporal gyrus and pSTS, extending up into overlying supramarginal gyrus in the right hemisphere only (Figs. 3*D*, 4*B*). Additional activations were observed in the right precentral gyrus and left superior frontal gyrus (Table 1).

Thus, shape discrimination enhanced activations in brain areas involved in processing stimulus features relevant to object processing (ventral stream), whereas discrimination of the rotational motion enhanced activations in V5/hMT+ and posterior superior temporal gyrus that have previously been implicated in visual and auditory motion processing (Zeki et al., 1991; Pavani et al., 2002; Warren et al., 2002; Poirier et al., 2005; Bartels et al., 2008; Sadaghiani et al., 2009; Scheef et al., 2009).

Interactions between synchrony and task

We next investigated where synchrony effects are not automatic but modulated by the specific task demands (i.e., shape vs motion discrimination). First, we tested for synchrony effects that were amplified for shape relative to motion discrimination [i.e., (Shape: SYNC > ASYN) > (Motion: SYNC > ASYN)]. This interaction contrast revealed activations in both left and right regions of lateral occipital cortex (Fig. 3*E*) that were in part spatially overlapping with the main effect of shape task and in part spatially overlapping with the main effect of synchrony, but rarely both (Fig. 4*A*). The parameter estimate plot from the peak voxel in the subcluster with coordinates corresponding with those of area LO in the right hemisphere (Malach et al., 1995) demonstrates a significantly greater enhancement in BOLD response for SYNC relative to ASYN trials when subjects discriminated shape

versus motion features (Fig. 3*E*, right). This interaction contrast also revealed activations in the superior/intraparietal sulcus and middle occipital gyrus (Table 1).

Second, we tested for synchrony effects that were amplified for motion relative to shape discrimination [i.e., (Motion: SYNC > ASYN) > (Shape: SYNC > ASYN)]. This interaction contrast revealed activations in the right pSTS (Fig. 3*F*) that overlapped almost entirely with the main effect of motion task and were located posterior to the main effect of synchrony (Fig. 4*B*). The parameter estimate plot from the interaction contrast's peak voxel in right pSTS illustrates that the increased response for synchronous relative to asynchronous audiovisual stimulation was greater during motion than shape discrimination (Fig. 3*F*, right).

Activations predicted by subjects' synchrony-induced behavioral benefit

To further investigate the behavioral relevance of synchrony-induced activation increases, we performed a regression analysis that used subjects' synchrony-induced behavioral benefit (observed only in the motion task) to predict their activation increases for synchronous relative to asynchronous trials. More specifically, we investigated whether subjects' behavioral benefit during the motion task (i.e., motion discrimination accuracy: SYNC > ASYN) predicted synchrony effects (i.e., motion discrimination activation: SYNC > ASYN) in the pSTS, as defined by significant voxels from the interaction between synchrony × motion and the anatomically defined bilateral V5/hMT+: the a priori region of interest given its well documented involvement in visual motion processing. Indeed, the synchrony effect

in right V5/hMT+ was significantly predicted by subjects' behavioral benefits in motion discrimination (corrected for multiple comparisons within a search mask that included left and right V5/hMT+, as defined by cytoarchitectonic probabilistic maps) (Fig. 5), but not within pSTS. The relationship between neural and behavioral synchrony effects suggests that response amplification within right V5/hMT+ may mediate subjects' improved motion discrimination for synchronous trials.

For further in-depth characterization (following a reviewer's suggestion), we investigated whether synchrony effects in any regions showing a main effect of synchrony, motion, or an interaction between synchrony and motion were also predicted by subjects' synchrony-induced behavioral benefit. Using a spherical search mask with a 10 mm radius centered on the peak voxel of each reported cluster (for coordinates, see Table 1), we found that only activation within the right anterior insula was significantly predicted by subjects' behavioral benefits in motion discrimination (corrected at $p < 0.05$ for multiple comparisons within the spherical search mask at the cluster level). These results suggest
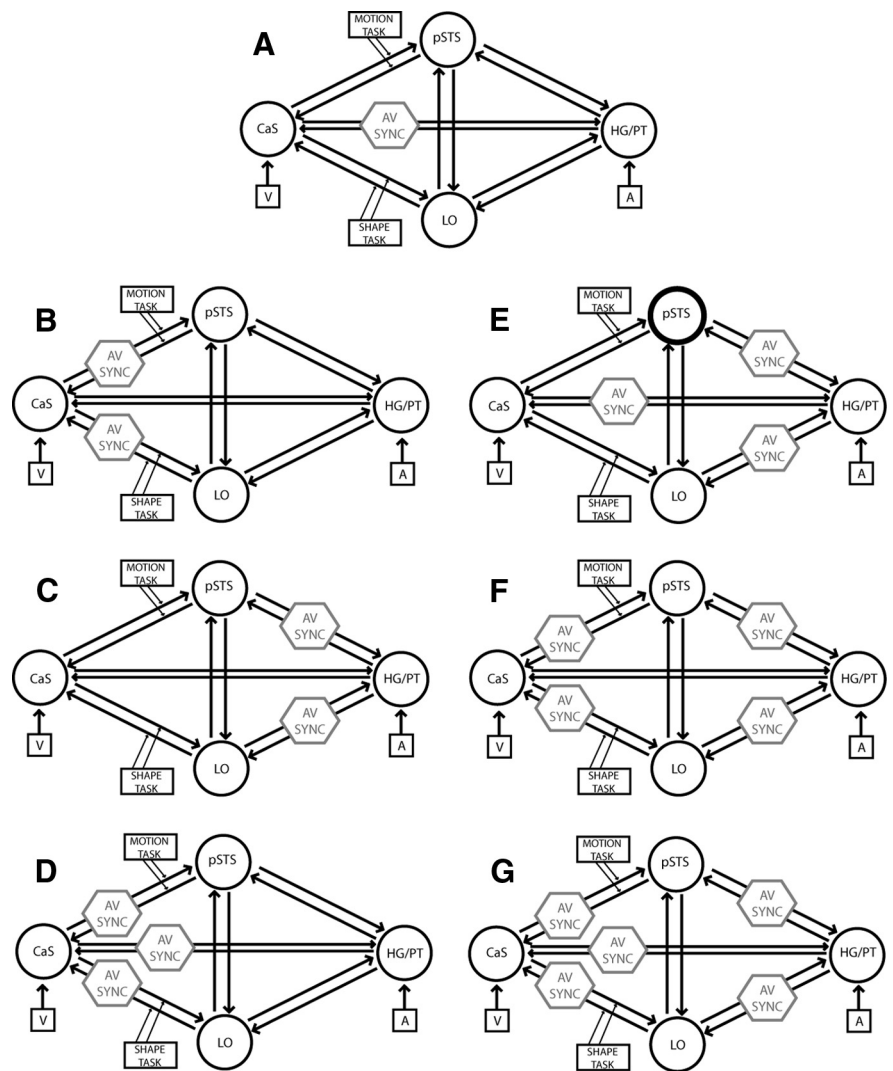

Figure 6. Seven candidate dynamic causal models were generated by manipulating the connection(s) that were modulated by audiovisual synchrony (AV SYNC). In all models, visual (V)/auditory (A) inputs entered into CaS and HG/PT, shape task modulated bidirectional connections between CaS and LO, and motion task modulated connections between CaS and posterior superior temporal sulcus.

that the right anterior insula may index how effectively audiovisual synchrony was detected (Bushara et al., 2001).

DCM results

Figure 6 shows the seven candidate dynamic causal models that differed with respect to the connection(s) that were modulated by audiovisual synchrony. Comparing model families allowing synchrony to modulate connectivity between CaS and HG/PT directly (A, D, E, G) and those that modulated connectivity indirectly (B, C, F) provided strong evidence for the direct family of models (posterior probability for direct family, >0.99). Figure 7 (inset) shows the results of the fixed-effects analysis, which provided strong evidence for model D with a model posterior probability of >0.94. As depicted in Figure 6, in this model, synchrony was allowed to modulate the connection strength of three bidirectional connections between CaS and HG/PT, CaS and the LO region, and CaS and the pSTS. In Figure 7, the numbers by the arrows representing the modulatory influences (i.e., tasks and synchrony) indicate the change in coupling (i.e., responsiveness of the target region to activity in the source region) induced by synchrony or task averaged across subjects. Synchrony signifi-
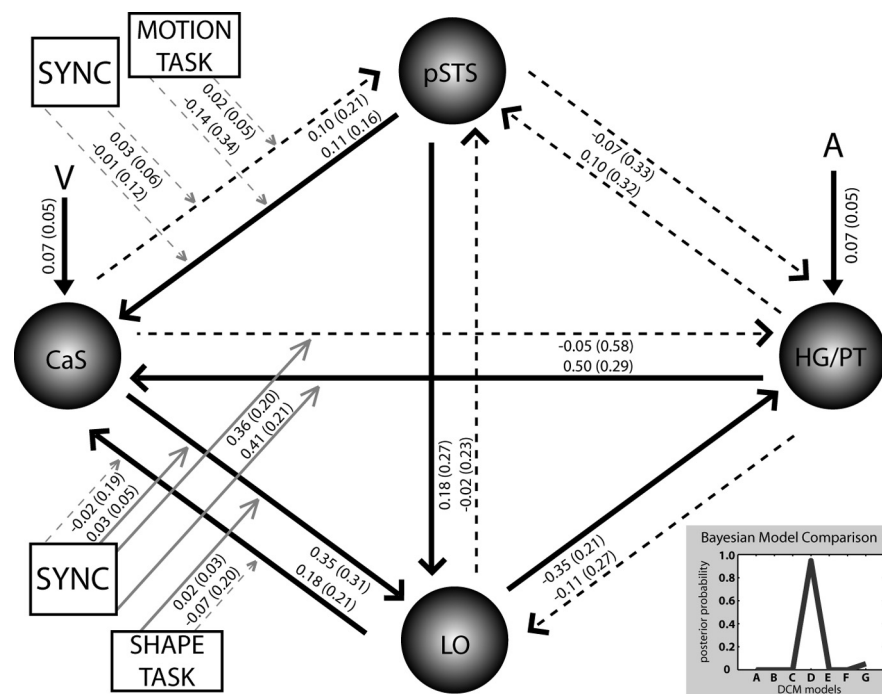
Figure 7. In the optimal model D (inset, posterior probability for each of the seven candidate DCMs), audiovisual stimulus synchrony (SYNC) enhances bidirectional connectivity between CaS and HG/PT and, to a lesser degree, from CaS to LO. Values are the across subject mean (SD in parentheses) of changes in connection strength (full lines significant at $p < 0.05$, dashed lines are nonsignificant). The bilinear parameters quantify how experimental manipulations change the values of intrinsic connections. A, Auditory; V, visual.

cantly ($p < 0.05$) enhanced bidirectional connectivity between CaS and HG/PT (Fig. 7). Synchrony also significantly enhanced forward connectivity from CaS to LO ($p < 0.05$). The shape task enhanced forward connectivity from CaS to LO ($p < 0.05$), yet there was only a nonsignificant trend toward enhanced forward connectivity during the motion task from CaS to pSTS ($p = 0.06$). Overall, the results of the DCM analysis suggest that the automatic response amplification during audiovisual synchrony occurs primarily via increased bidirectional effective connectivity between the calcarine sulcus and Heschl's gyrus/planum temporale and, to a lesser degree, through increased forward connectivity from calcarine sulcus to lateral occipital regions.

Discussion

Our results suggest that, under conditions of audiovisual synchrony, auditory stimuli may facilitate visual perception by amplifying responses in primary and higher-order visual and auditory cortices regardless of task-context, leading to enhanced stimulus salience. These synchrony effects may then propagate to shape-processing regions within lateral areas of occipital cortex/ventral parts of occipitotemporal cortex or pSTS, according to task requirements. Synchrony effects in right V5/hMT+ predicted subjects' improvements in motion discrimination for synchronous relative to asynchronous stimuli and may thus be behaviorally relevant.

Previous functional imaging studies investigating responses to nonspeech stimuli have revealed synchrony effects in a widely distributed neural system encompassing the superior colliculus, insula, occipital, inferior parietal, and superior/middle temporal regions. However, all of these studies have investigated synchrony effects for clearly discriminable sensory signals during passive exposure (Calvert et al., 2001; Dhamala et al., 2007, Noesselt et al., 2007) or in the context of explicit synchrony judgments

(Bushara et al., 2001). The present study is the first to investigate the role of synchrony in a more naturalistic behavioral context whereby the human brain actively extracts shape or motion information from partially obscured visual stimuli. Combining motion and shape discrimination in the same paradigm also enabled us to dissociate automatic from context-dependent synchrony effects.

In line with accumulating evidence implicating putatively sensory-specific and even primary sensory cortices in multisensory integration (Calvert et al., 1999; Schroeder and Foxe, 2002, 2005; Brosch et al., 2005; Ghazanfar et al., 2005; Ghazanfar and Schroeder, 2006; Kayser et al., 2007), we observed automatic synchrony effects in a widespread neural system encompassing bilateral CaS, inferior/middle occipital gyri, V5/hMT+, and HG/PT, regardless of task context. In the auditory domain, PT has been implicated generically in temporal pattern matching and sequencing (Griffiths and Warren, 2002). In audiovisual processing, PT may thus match temporal profiles of synchronous auditory and visual streams to assist initial scene segmentation. For instance, in the freezing phenomenon, coincident auditory cues facilitate segregation of a visual target from rapid, serially presented, visual distracter stimuli leading to improved detection of the visual target (Vroomen and de Gelder, 2000; Olivers and Van der Burg, 2008). In our experiment, response amplification in HG/PT and visual cortices may reflect successful temporal matching between the coincident auditory clicks and the partially obscured rotating patterns, leading to better segregation of the visual targets from the dynamic noise masks and consequently increased stimulus salience. One may speculate whether a behavioral benefit was observed only for motion discrimination (mean performance improvement, 6.2%) because motion processing places more demands on temporal sequencing than shape discrimination, which relies more on spatial processing.

Interestingly, the synchrony effects in higher order visual association areas were much more extensive than in previous passive exposure paradigms (Noesselt et al., 2007). This may result from subjects being actively engaged in shape/motion discrimination. Alternatively (or in addition), it may stem from degradation of our visual stimuli if the principle of inverse effectiveness (Meredith and Stein, 1983, 1986) not only governs multisensory enhancement per se, but also synchrony-induced response amplification. Thus, synchrony effects may also depend on stimulus efficacy and hence yield a more profound response enhancement for our visually obscured stimuli, becoming more salient when accompanied by coincident auditory clicks due to improved figure-ground segregation.

In addition to these general synchrony effects, ventral and lateral occipitotemporal regions, V5/hMT+, and pSTS also exhibited synchrony-induced activation increases that depended on the task context, i.e., whether motion or shape information was behaviorally relevant. Synchrony effects that were enhanced for shape discrimination were found in lateral occipital regions at voxel coordinates consistent with the functionally defined, classical,

visual-shape processing area LO (Malach et al., 1995; Vinberg and Grill-Spector, 2008). Previous multisensory studies have argued that LO responds to visual and tactile object inputs but not to object sounds as they cannot provide shape information (Amedi et al., 2002). In contrast, our study clearly demonstrates that coincident auditory clicks increase responses in the immediate vicinity of the shape area LO, particularly during shape discrimination, even though the clicks themselves provide no information about the shape of the visual patterns. Motion discrimination enhanced synchrony effects in motion area V5/hMT+ (Zeki et al., 1991; Bartels et al., 2008) and pSTS. A role in multisensory integration for pSTS is well established based on electrophysiological descriptions of bimodal neurons in nonhuman primates (Benevento et al., 1977; Bruce et al., 1981; Hikosaka et al., 1988) and more recently in several human brain imaging studies (Calvert et al., 2000, 2001; Beauchamp et al., 2004a,b; Stevenson et al., 2007; Stevenson and James, 2009; Werner and Noppeney, 2010a,b). Recent evidence also suggests a specific role for posterior regions of STS in processing second-order visual motion information (Noguchi et al., 2005). Furthermore, pSTS, together with auditory and visual regions, has recently been identified as being sensitive to temporal correspondence between simple transient auditory and visual events (Noesselt et al., 2007). Using two different task contexts, our study demonstrates that synchrony effects in pSTS are context sensitive and hence functionally dissociable from more automatic synchrony effects at low-level visual and auditory areas. In contrast to the well established role of pSTS in multisensory integration, V5/hMT+ has only recently been shown to be sensitive to auditory motion input (Poirier et al., 2005; Alink et al., 2008; Sadaghiani et al., 2009). Yet, to our knowledge, no one has previously demonstrated modulation of V5/hMT+ responses on the basis of temporal correspondence between stationary auditory stimuli and moving visual stimuli. In fact, investigating intersubject variability, we have demonstrated that synchrony effects in V5/hMT+ may be behaviorally relevant. Subjects' benefits in motion discrimination for synchronous relative to asynchronous trials significantly predicted their neural synchrony effects in V5/hMT+. These findings suggest that response amplification by coincident auditory clicks in V5/hMT+ may actually facilitate motion discrimination, as indexed by higher performance accuracy during motion discrimination. Future permanent or transient (e.g., transcranial magnetic stimulation) lesion studies are needed to corroborate the necessary role of V5/hMT+ in multisensory facilitation of motion discrimination. Collectively, the synchrony × task interactions in LO and pSTS suggest that the automatic synchrony effects in early visual cortex mediating increased target salience may be gated into higher functionally specialized areas according to task context: shape discrimination gates synchrony effects into shape processing area LO and motion discrimination into both pSTS and, when intersubject variation is taken into consideration, V5/hMT+.

Several neural mechanisms have been proposed to mediate multisensory interactions in sensory-specific areas, including feedforward thalamocortical, direct corticocortical, and feedback from classical multisensory areas (Schroeder and Foxe, 2005; Ghazanfar and Schroeder, 2006; Driver and Noesselt, 2008). We combined DCM and Bayesian model comparison to investigate how response amplification for synchronous stimuli emerges from distinct interactions among brain regions. All DCMs included the key regional players involved in our paradigm but differed in the connections that were modulated by synchrony (NB possibly attributable to methodological limits of fMRI, no thalamic effects were found and hence this potentially important locus could not be included). Bayesian model comparison strongly favored the model family that permitted synchrony to modulate direct connections between low-level visual and auditory areas over those that modulated connections between them indirectly. In the optimal DCM, audiovisual synchrony significantly enhanced connections between HG/PT and CaS bidirectionally and the forward connection from CaS to LO (NB direct effective connectivity does not implicate direct anatomical connectivity). Collectively, the DCM analysis suggests that audiovisual synchrony amplifies responses in CaS and HG/PT via direct, mutually excitatory interactions between the two regions. These context-independent synchrony effects then propagate to LO or pSTS according to the behavioral relevance of shape and motion features through context-sensitive modulation of the forward connections from CaS to LO or to pSTS, respectively.

In conclusion, this study dissociated three mechanisms by which synchrony can modulate the neural processing of audiovisual signals. First, in HG/PT and primary visual cortex (CaS), synchrony amplified audiovisual processing, regardless of task context, via enhanced direct effective connectivity. Second, in bilateral regions of LO cortex, synchrony effects were enhanced for shape discrimination and, in the right pSTS, for motion discrimination. Third, in right V5/hMT+, synchrony effects were predicted by subject-specific, synchrony-induced, behavioral benefit in motion discrimination. To increase the salience of visually obscured moving objects, audiovisual synchrony may amplify visual activations by increasing the excitatory interactions between CaS and HG/PT. The enhanced CaS responses for synchronous stimuli may then be gated to the relevant functionally specialized areas (e.g., shape area LO) via contextually increased forward connectivity from CaS to LO/pSTS.

References

Alink A, Singer W, Muckli L (2008) Capture of auditory motion by vision is represented by an activation shift from auditory to visual motion cortex. J Neurosci 28:2690–2697.

Amedi A, Jacobson G, Hendler T, Malach R, Zohary E (2002) Convergence of visual and tactile shape processing in the human lateral occipital complex. Cereb Cortex 12:1202–1212.

Bartels A, Logothetis NK, Moutoussis K (2008) fMRI and its interpretations: an illustration on directional selectivity in area V5/MT. Trends Neurosci 31:444–453.

Beauchamp MS, Lee KE, Argall BD, Martin A (2004a) Integration of auditory and visual information about objects in superior temporal sulcus. Neuron 41:809–823.

Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A (2004b) Unraveling multisensory integration: patchy organization within human STS multisensory cortex. Nat Neurosci 7:1190–1192.

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.

Brosch M, Selezneva E, Scheich H (2005) Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. J Neurosci 25:6797–6806.

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.

Bushara KO, Grafman J, Hallett M (2001) Neural correlates of auditory-visual stimulus onset asynchrony detection. J Neurosci 21:300–304.

Calvert GA, Brammer MJ, Bullmore ET, Campbell R, Iversen SD, David AS (1999) Response amplification in sensory-specific cortices during crossmodal binding. Neuroreport 10:2619–2623.

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, Hansen PC, Iversen SD, Brammer MJ (2001) Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. Neuroimage 14:427–438.

Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE (1991) Selective and divided attention during visual discriminations of shape, colour and speed: functional anatomy by positron emission tomography. J Neurosci 11:2383–2402.

Dhamala M, Assisi CG, Jirsa VK, Steinberg FL, Kelso JA (2007) Multisensory integration for timing engages different brain networks. Neuroimage 34:764–773.

Driver J, Noesselt T (2008) Multisensory interplay reveals crossmodal influences on "sensory-specific" brain regions, neural responses, and judgments. Neuron 57:11–23.

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25:1325–1335.

Foxe JJ, Schroeder CE (2005) The case for feedforward multisensory convergence during early cortical processing. Neuroreport 16:419–423.

Frassinetti F, Bolognini N, Làdavas E (2002) Enhancement of visual perception by crossmodal visuo-auditory interaction. Exp Brain Res 147:332–343.

Friston KJ, Penny WD, Phillips C, Kiebel SJ, Hinton G, Ashburner J (2002) Classical and Bayesian inference in neuroimaging: theory. NeuroImage 16:465–483.

Friston K (2009) Dynamic causal modeling and Granger causality. Comments on: The identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution. Neuroimage. Advance online publication. Retrieved Aug. 2, 2010. doi:10.1016/j.neuroimage.2009.09.031.

Friston KJ, Holmes AP, Worsley KJ (1999) How many subjects constitute a study? Neuroimage 10:1–5.

Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. Neuroimage 19:1273–1302.

Ghazanfar AA, Schroeder CE (2006) Is neocortex essentially multisensory? Trends Cogn Sci 10:278–285.

Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK (2005) Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. J Neurosci 25:5004–5012.

Griffiths TD, Warren JD (2002) The planum temporale as a computational hub. Trends Neurosci 25:348–353.

Heron J, Whitaker D, McGraw PV (2004) Sensory uncertainty governs the extent of audio-visual interaction. Vision Res 44:2875–2884.

Hikosaka K, Iwai E, Saito H, Tanaka K (1988) Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. J Neurophysiol 60:1615–1637.

Kayser C, Logothetis NK (2007) Do early sensory cortices integrate crossmodal information? Brain Struct Funct 212:121–132.

Kayser C, Petkov CI, Augath M, Logothetis NK (2007) Functional imaging reveals visual modulation of specific fields in auditory cortex. J Neurosci 27:1824–1835.

Kiebel SJ, Klöppel S, Weiskopf N, Friston KJ (2007) Dynamic causal modeling: a generative model of slice timing in fMRI. Neuroimage 34: 1487–1496.

Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc Natl Acad Sci U S A 92:8135–8139.

Malikovic A, Amunts K, Schleicher A, Mohlberg H, Eickhoff SB, Wilms M, Palomero-Gallagher N, Armstrong E, Zilles K (2007) Cytoarchitectonic analysis of the human extrastriate cortex in the region of V5/MT+: a probabilistic, stereotaxic map of area hOc5. Cereb Cortex 17:562–574.

Meredith MA, Stein BE (1983) Interactions among converging sensory inputs in the superior colliculus. Science 221:389–391.

Meredith MA, Stein BE (1986) Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. J Neurophsyiol 56:640–662.

Munhall KG, Gribble P, Sacco L, Ward M (1996) Temporal constraints on the McGurk effect. Percept Psychophys 58:351–362.

Noesselt T, Rieger JW, Schoenfeld MA, Kanowski M, Hinrichs H, Heinze HJ, Driver J (2007) Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. J Neurosci 27:11431–11441.

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

Noguchi Y, Kaneoke Y, Kakigi R, Tanabe HC, Sadato N (2005) Role of the superior temporal region in human visual motion perception. Cereb Cortex 15:1592–1601.

Olivers CN, Van der Burg E (2008) Bleeping you out of the blink: sound saves vision from oblivion. Brain Res 1242:191–199.

Pandey PC, Kunov H, Abel SM (1986) Disruptive effects of auditory signal delay on speech perception with lipreading. J Aud Res 26:27–41.

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

Penny WD, Stephan KE, Mechelli A, Friston KJ (2004a) Comparing dynamic causal models. Neuroimage 22:1157–1172.

Penny WD, Stephan KE, Mechelli A, Friston KJ (2004b) Modelling functional integration: a comparison of structural equation and dynamic causal models. Neuroimage 23 [Suppl 1]:S264–S274.

Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston K, Schofield TM, Leff AP (2010) Comparing families of dynamic causal models. PLoS Comput Biol 6:e1000709.

Peuskens H, Claeys KG, Todd JT, Norman JF, Van Hecke P, Orban GA (2004) Attention to 3-D shape, 3-D motion and texture in 3-D structure from motion displays. J Cogn Neurosci 16:665–682.

Poirier C, Collignon O, Devolder AG, Renier L, Vanlierde A, Tranduy D, Scheiber C (2005) Specific activation of the V5 brain area by auditory motion processing: an fMRI study. Brain Res Cogn Brain Res 25:650–658.

Roebroeck A, Formisano E, Goebel R (2009) The identification of interacting networks in the brain using fMRI: Model selection, causality and deconvolution. Neuroimage. In press.

Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ (2007) Do you see what I am saying?: exploring visual enhancement of speech comprehension in noisy environments. Cereb Cor 17:1147–1153.

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

Sarkheil P, Vuong QC, Bülthoff HH, Noppeney U (2008) The integration of higher order form and motion by the human brain. Neuroimage 42:1529–1536.

Scheef L, Boecker H, Daamen M, Fehse U, Landsberg MW, Granath DO, Mechling H, Effenberg AO (2009) Multimodal motion processing in area V5/MT: evidence from an artificial class of audio-visual events. Brain Res 1252:94–104.

Schroeder CE, Foxe J (2005) Multisensory contributions to low-level, "unisensory" processing. Curr Opin Neurobiol 15:454–458.

Schroeder CE, Foxe JJ (2002) The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. Brain Res Cogn Brain Res 14:187–198.

Sekuler R, Sekuler AB, Lau R (1997) Sound alters visual motion perception. Nature 385:308.

Shams L, Kamitani Y, Shimojo S (2002) Visual illusion induced by sound. Brain Res Cogn Brain Res 14:147–152.

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

Stevenson RA, James TW (2009) Audiovisual integration in human superior temporal sulcus: inverse effectiveness and the neural processing of speech and object recognition. Neuroimage 44:1210–1223.

Stevenson RA, Geoghegan ML, James TW (2007) Superadditive BOLD activation in superior temporal sulcus with threshold non-speech objects. Exp Brain Res 179:85–95.

Vinberg J, Grill-Spector K (2008) Representation of shapes, edges, and surfaces across multiple cues in the human visual cortex. J Neurophysiol 99:1380–1393.

Vroomen J, de Gelder B (2000) Sound enhances visual perception: crossmodal effects of auditory organization on vision. J Exp Psychol Hum Percept Perform 26:1583–1590.

Warren JD, Zielinski BA, Green GG, Rauschecker JP, Griffiths TD (2002) Perception of sound-source motion by the human brain. Neuron 34:139–148.

Werner S, Noppeney U (2010a) Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. Cereb Cortex 20:1829–1842.

Werner S, Noppeney U (2010b) Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. J Neurosci 30:2662–2675.

Zeki S, Watson JD, Lueck CJ, Friston KJ, Kennard C, Frackowiak RS (1991) A direct demonstration of functional specialization in the human visual cortex. J Neurosci 11:641–649.