Behavioral/Systems/Cognitive

Audiovisual Temporal Correspondence Modulates Human Multisensory Superior Temporal Sulcus Plus Primary Sensory Cortices

Toemme Noesselt,[1,2] **Jochem W. Rieger,**[2] **Mircea Ariel Schoenfeld,**[2] **Martin Kanowski,**[2] **Hermann Hinrichs,**[2] **Hans-Jochen Heinze,**[2] **and Jon Driver**[1]

[1]UCL Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, United Kingdom, and [2]Department of Neurology II and Center for Advanced Imaging, Otto-von-Guericke-Universität, 39120 Magdeburg, Germany

The brain should integrate related but not unrelated information from different senses. Temporal patterning of inputs to different modalities may provide critical information about whether those inputs are related or not. We studied effects of temporal correspondence between auditory and visual streams on human brain activity with functional magnetic resonance imaging (fMRI). Streams of visual flashes with irregularly jittered, arrhythmic timing could appear on right or left, with or without a stream of auditory tones that coincided perfectly when present (highly unlikely by chance), were noncoincident with vision (different erratic, arrhythmic pattern with same temporal statistics), or an auditory stream appeared alone. fMRI revealed blood oxygenation level-dependent (BOLD) increases in multisensory superior temporal sulcus (mSTS), contralateral to a visual stream when coincident with an auditory stream, and BOLD decreases for noncoincidence relative to unisensory baselines. Contralateral primary visual cortex and auditory cortex were also affected by audiovisual temporal correspondence or noncorrespondence, as confirmed in individuals. Connectivity analyses indicated enhanced influence from mSTS on primary sensory areas, rather than vice versa, during audiovisual correspondence. Temporal correspondence between auditory and visual streams affects a network of both multisensory (mSTS) and sensory-specific areas in humans, including even primary visual and auditory cortex, with stronger responses for corresponding and thus related audiovisual inputs.

Key words: audiovisual; temporal integration; connectivity; fMRI; human multisensory

Introduction

Signals entering different senses can sometimes originate from the same object or event. The brain should integrate just those multisensory inputs that reflect a common external source, as may be indicated by spatial, temporal, or semantic constraints (Stein and Meredith, 1993; Calvert et al., 2004; Spence and Driver, 2004; Macaluso and Driver, 2005; Schroeder and Foxe, 2005; Ghazanfar and Schroeder, 2006). Many neuroscience and human neuroimaging studies have investigated possible "spatial" constraints on multisensory integration (Wallace et al., 1996; Macaluso et al., 2000, 2004; McDonald et al., 2000, 2003), or factors that may be more "semantic," such as matching vocal sounds and mouth movements (Calvert et al., 1997; Ghazanfar et al., 2005), or visual objects that match environmental sounds (Beauchamp et al., 2004a,b; Beauchamp, 2005a).

Here we focus on possible constraints from "temporal" correspondence only (Stein et al., 1993; Calvert et al., 2001; Bischoff et al., 2007; Dhamala et al., 2007), using streams of nonsemantic stimuli (visual transients and beeps) to isolate purely temporal influences. We arranged that audiovisual temporal relations conveyed strong information that auditory and visual streams were related or unrelated, by using erratic, arrhythmic temporal patterns that either matched perfectly between audition and vision (very unlikely by chance) or mismatched substantially but had the same overall temporal statistics. We anticipated increased brain activations for temporally corresponding audiovisual streams (compared with noncorresponding or unisensory) in multisensory superior temporal sulcus (mSTS). This region receives converging auditory and visual inputs (Kaas and Collins, 2004) and is thought to contribute to multisensory integration (Benevento et al., 1977; Bruce et al., 1981; Cusick, 1997; Beauchamp et al., 2004b). mSTS was influenced by audiovisual synchrony in some previous function magnetic resonance imaging (fMRI) studies that used very different designs and/or more semantic stimuli (Calvert et al., 2001; van Atteveldt et al., 2006; Bischoff et al., 2007; Dhamala et al., 2007).

There have been several recent proposals that multisensory interactions may affect not only established multisensory brain regions (such as mSTS) but also brain areas (or evoked responses) traditionally considered sensory specific (for review, see

Brosch and Scheich, 2005; Foxe and Schroeder, 2005; Ghazanfar and Schroeder, 2006), although some event-related potential (ERP) examples proved controversial (Teder-Sälejärvi et al., 2002). Given recent results from invasive recording in monkey primary auditory cortex (Brosch et al., 2005; Ghazanfar et al., 2005; Lakatos et al., 2007), we anticipated that audiovisual correspondence in temporal patterning might affect sensory-specific "auditory" cortex. We tested this with human whole-brain fMRI, which also allowed assessment of any impact on sensory-specific "visual" cortex (and mSTS) concurrently. Finally, we assessed effective connectivity (or functional coupling) between the areas that were differentially activated by audiovisual temporal correspondence (AVC) [vs noncorrespondence (NC)]. We found that audiovisual correspondence in temporal patterning can affect both primary visual and auditory cortex in humans, as well as mSTS, with some evidence for feedback influences from mSTS in our paradigm.

Materials and Methods

Twenty-four neurologically normal subjects (10 female; mean age, 24 years) participated after informed consent in accord with local ethics. Visual stimulation was in the upper left hemifield for 12 subjects and in the upper right for the other 12. This was presented at the top of the MR bore via clusters of four optic fibers arranged into a rectangular shape and five interleaved fibers arranged into a cross shape, 2° above the horizontal meridian at an eccentricity of 18°. Visual stimuli were presented peripherally, which may maximize the opportunity for interplay between auditory and visual cortex (Falchier et al., 2002), and also allowed us to test for any contralaterality in effects for one visual field or the other. The peripheral fiber-optic endings could be illuminated red or green with a standard luminance of 40 cd/m^2 and were 1.5° in diameter (for schematics of the resulting colored "shapes," see Fig. 1C). Streams of visual transients were produced by switching between the differently colored cross and rectangle shapes (red and green, respectively, in Fig. 1C, but shape-color was counterbalanced across subjects). Throughout each experimental run, subjects fixated a central fixation cross of ~0.2° in diameter. Eight red–green (cross/square) reversals occurred in a 2 s interval, with the stimulus-onset asynchrony (SOA) between each successive color change ranging in a pseudorandom manner from 100 to 500 ms (mean reversal rate of 4 Hz, with rectangular distribution from 2 to 10 Hz, but note that reversal rate was never constant for successive transients), to produce a uniquely jittered, highly arrhythmic timing for each stream.

Auditory stimuli were presented via a piezoelectric speaker inside the scanner, just above fixation. Each auditory stimulus was a clearly audible 1 kHz sound burst with duration of 10 ms at ~70 dB. Identical temporally jittered stimulation sequences within vision and/or audition were used in all conditions overall (fully counterbalanced), so that there was no difference whatsoever in temporal statistics between conditions, except for the critical temporal relationship between auditory and visual streams during multisensory trials (unisensory conditions were also included; see below). Fourier analyses of the amplitude spectra for all the stimulus trains used indicated that no frequency was particularly prominent across a range of 1.5–20 Hz.

The experimental stimuli (for the visual-only baseline, auditory-only baseline, and for AVC or NC) were all presented during silent periods (2 s) interleaved with scanning (3 s periods of fMRI acquisition) to prevent scanner noise interfering with our auditory stimuli or perception of their temporal relationship with visual flashes. In the AVC condition, a tone burst was initiated synchronously with every visual transient (Fig. 1A) and thus had exactly the same jittered, arrhythmic temporal pattern. During the NC condition (Fig. 1B), tone bursts occurred with a different temporal pattern (but always having the same overall temporal statistics, including mean rate of 4 Hz within a rectangular distribution from 2 to 10 Hz, and a highly arrhythmic nature), with a minimal protective "window" of 100 ms now separating each sound from onset of a visual pattern reversal (Fig. 1B).

This provided clear information that the two streams were either strongly related, as in the AVC condition (such perfect coincidence for the erratic, arrhythmic temporal patterns is exceptionally unlikely to arise by chance), or were unrelated, as for the NC condition. During the latter noncoincidence, up to two events in one stream could occur before an event in the second stream had to occur. The mean 4 Hz stimulation rate (range of 2–10) used here, together with the constraints (protective window, see Fig. 1B) implemented to avoid any accidental synchronies in the noncorresponding condition, should optimize detection of audiovisual correspondence versus noncorrespondence (Fujisaki et al., 2006) but make these bimodal conditions otherwise identical in terms of the temporal patterns presented overall to each modality. All sequences were created individually for each subject using Matlab 6.5 (MathWorks, Natick, MA). Piloting confirmed that the correspondence versus noncorrespondence relationship could be discriminated readily when requested (mean percentage correct, 93.8%), even with such peripheral visual stimuli. Irregular stimulus trains [rather than rhythmic (cf. Lakatos et al., 2005)] were chosen, because this makes an audiovisual temporal relationship much less likely to arise by chance alone, and hence (a)sychrony typically becomes easier to detect than for regular frequencies or for single auditory and visual events rather than stimulus trains (Slutsky and Recanzone, 2001; Noesselt et al., 2005).

Two "unisensory" conditions (i.e., visual or auditory streams alone) were also run. These allowed our fMRI analysis to distinguish candidate "multisensory" brain regions (responding to either type of unisensory stream) from sensory-specific regions (visually or auditorily selective; see below).

Throughout each experimental run, participants performed a central visual monitoring task requiring detection of occasional brief (1 ms) brightening of the fixation point via button press. This could occur at random times (average rate of 0.1 Hz) during both stimulation and scan periods. Participants were instructed to perform this fixation-monitoring task, and auditory and peripheral visual stimuli were always task irrelevant. We chose this fixation-monitoring task to avoid the different multisensory conditions being associated with changes in performance that might otherwise have contaminated the fMRI data because we were interested in stimulus-determined (rather than task-determined) effects of audiovisual temporal correspondence and also so as to minimize eye movements. Eye position was monitored on-line during scanning (Kanowski et al., 2007).

fMRI data were collected in four runs with a neuro-optimized 1.5 GE (Milwaukee, WI) scanner equipped with a head–spine coil. A rapid sparse-sampling protocol was used (136 volumes per run with 30 slices covering whole brain; repetition time of 3 s; silent pause of 2 s; echo time of 40 ms; flip angle of 90°; resolution of 3.5 × 3.5 mm; 4 mm slice thickness; field of view was 20 cm). Experimental stimuli were presented during the silent scanner periods (2 s scanner pauses). Each mini-block lasted 20 s per condition, containing 8 s (4 × 2) of stimulation (with each successive 2 s segment of stimuli then separated by 3 s of scanning). These mini-blocks of experimental stimulation in one of the four conditions or another (random sequence) were each separated by 20 s blocks, in which only the central fixation task was presented (unstimulated blocks).

After preprocessing for motion correction, normalization, and 6 mm smoothing, data were analyzed in SPM2 (Wellcome Department of Cognitive Neurology, University College London, London, UK) by modeling the four conditions and the intervening unstimulated baselines with box-car functions. Voxel-based group effects were assessed with a second-level random-effects analysis, identifying candidate multisensory regions (responding to both auditory and visual stimulation), sensory-specific regions (difference between visual minus auditory or vice versa), and the critical differential effects of coincident minus noncoincident audiovisual presentations.

Conjunction analyses assessed activation within sensory-specific and multisensory cortex (thresholded at $p < 0.001$), within areas that also showed a significant modulation of the omnibus F test at $p < 0.001$ (Beauchamp, 2005b) for clusters of >20 contiguous voxels. To confirm localization to a particular anatomical region (e.g., calcarine sulcus) in

individuals, we extracted beta estimates of blood oxygenation level-dependent (BOLD) modulation for each condition, from their local maxima for the comparison AVC > NC, within regions of interest (ROIs) comprising early visual and auditory cortex and within mSTS. These ROIs were initially identified via a combination of anatomical criteria (calcarine sulcus, medial part of anteriormost Heschl's gyrus, posterior STS) and functional criteria in each individual (i.e., sensory-specific responses to our visual or auditory stimuli, for calcarine sulcus or Heschl's gyrus, respectively, or multisensory response to both modalities in the case of mSTS). We then tested the voxels within these individually defined ROIs for any impact of the critical manipulation (which was orthogonal to the contrasts identifying those ROIs) of audiovisual correspondence minus noncorrespondence. We also compared each of those two multisensory conditions with the unimodal baselines for the same regions on the extracted data.

Finally, we used connectivity analyses to assess possible influences (or "functional coupling") between affected mSTS, primary visual cortex (V1), and primary auditory cortex (A1) for the fMRI data. We first used the established "psychophysiological interaction" (PPI) (Friston et al., 1997) approach, which is relatively assumption free. This assesses condition-specific covariation between a seeded brain area and any other regions, for the residual variance that remains after mean BOLD effects attributable to condition have been discounted. Data from the left visual field (LVF) group were left–right flipped to allow pooling with the right visual field (RVF) group for this and to assess any effects that generalized across hemispheres (Lipschutz et al., 2002). PPI analyses can serve to establish condition-dependent functional coupling (or "effective connectivity") between brain regions but do not provide information about the predominant direction of influence of information transfer. Accordingly, we further assessed potential influences between mSTS, V1, and A1 with a directed information transfer (DIT) measure, as developed recently (Hinrichs et al., 2006). DIT assesses predictability of one time series from another, in a data-driven approach that makes minimal assumptions. If the joint time series for, say, regions A and B predict future signals in time series B, better than B does alone, this is taken to indicate that A influences B with a strength indicated by the corresponding DIT measure. If DIT from A to B is larger than vice versa, this indicates directed information flow from A to B. Our DIT analysis used 96 time points (four runs of four blocks with six points per block) per condition and region. From these data, we derived the DIT values from the current samples of A and B to the subsequent sample of B, and vice versa, and then averaged over all 96 samples. Here we used the DIT approach to assess possible pairwise relationships between mSTS, V1, and A1 for their extracted time series, assessing DIT measures for all pairings between these (i.e., V1–A1, V1–STS, or A1–STS) with paired t tests.

Results

Subjects performed the monitoring task on the fixation point (Fig. 1) (see Materials and Methods) equally well (mean 83% accuracy) in all conditions (all $p > 0.2$), with maintenance of central fixation also equally good across conditions [i.e., similar performance for all conditions (<2° deviation in 98% of trials)], as expected given the task at central fixation.

Modulation of BOLD responses attributable to audiovisual correspondence

For fMRI analyses, the random-effect SPM analysis confirmed that unisensory visual streams activated sensory-specific occipital visual cortex, as expected, whereas auditory streams activated auditory core, belt, and parabelt regions in temporal cortex, as expected (Table 1). Candidate multisensory regions, activated by both the unisensory visual and unisensory auditory streams, included bilateral posterior STS and posterior parietal and dorsolateral prefrontal areas. However, within these candidate multisensory regions, only STS showed the critical effects of
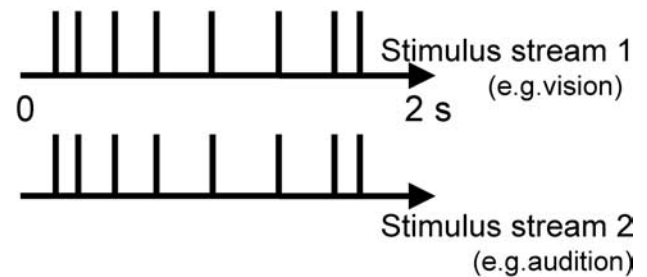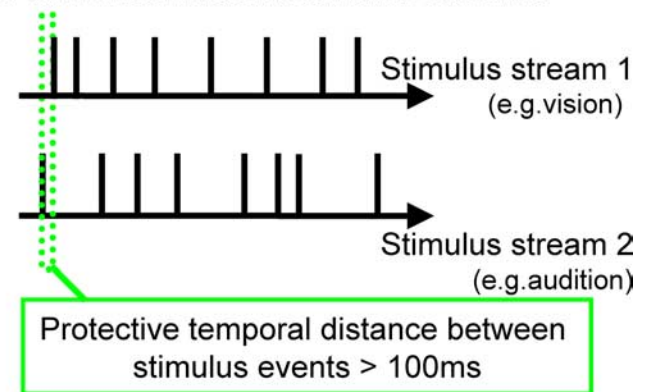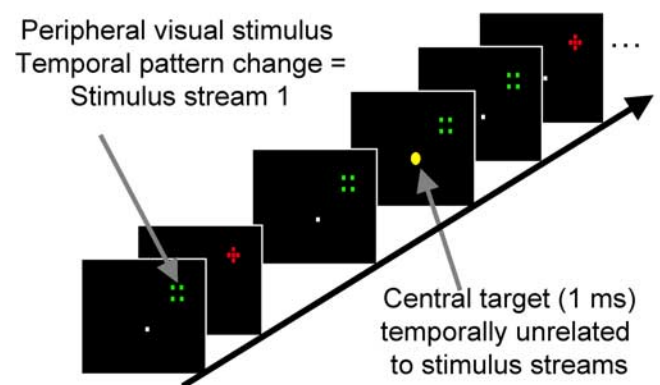

Figure 1. Schematic illustration of stimulus sequences and setup. ***A***, Illustrative examples of timing for sequences in vision (top row) and, in audition, for the audiovisual correspondence condition (i.e., perfectly synchronous sequence, with jittered arrhythmic timing, average rate of 4 Hz, and rectangular distribution of 2–10 Hz). ***B***, This example illustrates the noncorresponding condition; the two streams still have comparable stimulus rate (and other temporal statistics) overall but are now highly unrelated (differently jittered arrhythmic sequences, with a protective minimal window of 100 ms separating visual and auditory onsets; see green dotted lines). ***C***, Example visual stimuli are depicted. Participants maintained central fixation, whereas optic fibers at 18° eccentricity were illuminated to produce a red cross stimulus or a nonoverlapping green square stimulus, with successive alternation between these. The task was to monitor the central yellow fixation light-emitting diode for occasional brightening (indicated here by enlarged central yellow dot; duration of 1 ms and average occurrence of 0.1 Hz), with timing unrelated to the task-irrelevant auditory or visual streams.

audiovisual temporal correspondence (Table 2A, Fig. 2A). Within the functionally defined multisensory regions, AVC minus NC specifically activated (at $p < 0.001$) the contralateral mSTS (i.e., right mSTS for LVF group, peak at 60, −48, 12; left

Table 1. BOLD effects in group average

	MNI coordinates				
	Peak				
Brain region	t	p <	x	y	z
Multisensory areas: visual condition > baseline and auditory condition > baseline[a]					
Visual stimuli in RVF					
Contralateral left hemisphere					
Posterior STS	4.14	0.001	−66	−48	12
Superior parietal lobe	3.64	0.001	−24	−60	38
Middle frontal gyrus	4.54	0.001	−40	8	26
Ipsilateral right hemisphere					
Posterior STS	2.35	0.009	−58	−44	12
Superior parietal lobe	3.53	0.001	−34	−48	46
Middle frontal gyrus	3.76	0.001	−40	20	18
Visual stimuli in LVF					
Contralateral right hemisphere					
Posterior STS	3.92	0.001	62	−44	8
Superior parietal lobe	2.60	0.005	24	−64	42
Middle frontal gyrus	3.14	0.001	44	14	34
Ipsilateral left hemisphere					
Posterior STS	3.30	0.001	−62	−54	18
Superior parietal lobe	2.58	0.005	−20	−56	54
Middle frontal gyrus	3.46	0.001	−40	16	16
Visual areas: unimodal visual > auditory condition[b]					
Visual stimuli in RVF					
Contralateral left hemisphere					
Lingual gyrus	7.50	0.001	−12	−76	−4
Fusiform gyrus	7.30	0.001	−18	−72	−8
Fusiform gyrus	5.31	0.001	−34	−66	−18
Transversal occipital sulcus	4.38	0.001	−20	−90	32
Middle occipital sulcus	3.95	0.001	−44	−78	10
Superior occipital gyrus	4.65	0.001	−32	−80	18
Visual stimuli in LVF					
Contralateral right hemisphere					
Lingual gyrus/calcarine	7.29	0.001	6	−76	0
Fusiform gyrus	7.31	0.001	14	−72	−10
Fusiform gyrus	4.90	0.001	26	−66	−6
Transversal occipital sulcus	6.74	0.001	26	−86	26
Middle occipital sulcus	5.03	0.001	42	−70	8
Superior occipital gyrus	5.69	0.001	30	−82	20
Posterior parietal lobe	3.94	0.001	22	−70	60
Auditory areas: unimodal auditory > visual condition[c]					
Visual stimuli in RVF					
Left hemisphere					
Posterior Heschl's gyrus/insula	4.64	0.001	−40	−26	4
Middle STS	5.27	0.001	−48	−32	6
Middle STS	5.27	0.001	−48	−32	6
Middle STS	5.03	0.001	−48	−14	−6
Right hemisphere					
Planum temporale	3.89	0.001	56	−26	4
Insula/planum polare	4.19	0.001	52	0	−8
Middle STS	3.89	0.001	62	−30	6
Visual stimuli in LVF					
Right hemisphere					
Planum temporale	4.07	0	56	−24	4
Insula	3.81	0	40	−20	−2
Middle STS	3.58	0	62	−34	2
Left hemisphere					
Posterior Heschl's gyrus/PT	3.83	0	−48	−22	10
Middle STS	4.39	0	−44	−32	6
Middle STS	3.91	0	−54	−20	4

Group average activation peaks for the contrasts testing for candidate multisensory regions (i.e., responding to both visual and auditory unimodal conditions relative to unstimulated baseline) or sensory-specific regions (i.e., significant effect of visual minus auditory stimulation or vice versa). Only clusters containing >20 voxels are described (see Materials and Methods). MNI, Montreal Neurological Institute.

[a]Candidate multisensory areas: conjunction of unimodal auditory condition minus baseline and unimodal visual condition minus baseline.

[b]Visual sensory-specific cortex: unimodal visual minus unimodal auditory conditions.

[c]Auditory sensory-specific cortex: unimodal auditory minus unimodal visual conditions.

Table 2. Group average activation peaks for the experimental contrast audiovisual coincidence > noncoincidence within multisensory or sensory-specific regions (i.e., significant effect of visual minus auditory stimulation or vice versa)

	MNI coordinates				
Brain region	Peak t	p <	x	y	z
Multisensory areas: audiovisual coincident > noncoincident condition					
Visual stimuli in RVF					
Contralateral left hemisphere					
Posterior STS	3.43	0.001	−54	−50	8
Visual stimuli in LVF					
Contralateral right hemisphere					
Posterior STS	3.19	0.001	60	−48	12
Visual areas: audiovisual coincident > noncoincident condition					
Visual stimuli in RVF					
Contralateral left hemisphere					
Lingual gyrus/					
calcarine	3.01	0.001	−8	−76	2
Fusiform gyrus	4.70	0.001	−30	−74	−16
Transversal occipital					
sulcus	4.38	0.001	−26	−88	30
Visual stimuli in LVF					
Contralateral right hemisphere					
Calcarine	3.07	0.001	12	−82	10
Fusiform gyrus	4.70	0.001	24	−74	−4
Transversal occipital					
sulcus	5.14	0.001	22	−88	36
Auditory areas: audiovisual coincident > noncoincident condition					
Visual stimuli in RVF					
Left hemisphere					
Heschl's gyrus	4.08	0.001	−50	−16	8
Middle STG	3.03	0.001	−66	−28	12
Planum polare	2.95	0.005	−48	−14	−6
Right hemisphere					
Planum polare	3.89	0.001	54	−4	−8
Middle STG	4.19	0.001	56	−18	2
Middle STS	3.89	0.001	50	−10	−4
Visual stimuli in LVF					
Right hemisphere					
Heschl's gyrus	3.46	0.001	48	−20	10
Planum temporale/					
STS	3.46	0.001	54	−28	6
Planum polare	3.86	0.001	−40	−22	−4
Left hemisphere					
Planum temporale	4.18	0.001	38	−36	18
Middle STG	3.55	0.001	64	−34	14

Only clusters containing >20 voxels are described (see Materials and Methods and Fig. 2). MNI, Montreal Neurological Institute; STG, superior temporal gyrus.

mSTS for RVF group, peak at −54, −52, 8) (Fig. 2A). Additional tests on individually defined maxima within mSTS confirmed that, contralateral to the visual stream, responses to AVC were significantly elevated not only relative to the NC condition but also relative to either unisensory stream alone ($p < 0.03$). Noncoincidence led instead to a reliably decreased response relative to either unisensory baseline ($p < 0.01$) (see bar graph for mSTS in Fig. 2A). All individual subjects showed this pattern (for illustrative single subject, see Fig. 3A).

Importantly, an analogous pattern was found within sensory-specific cortices. For visual cortex, we found increased BOLD responses for the AVC > NC comparison near the contralateral calcarine fissure (peaks at −12, −76, 0 and 12, −82, 12 for RVF and LVF groups, respectively; both $p <$
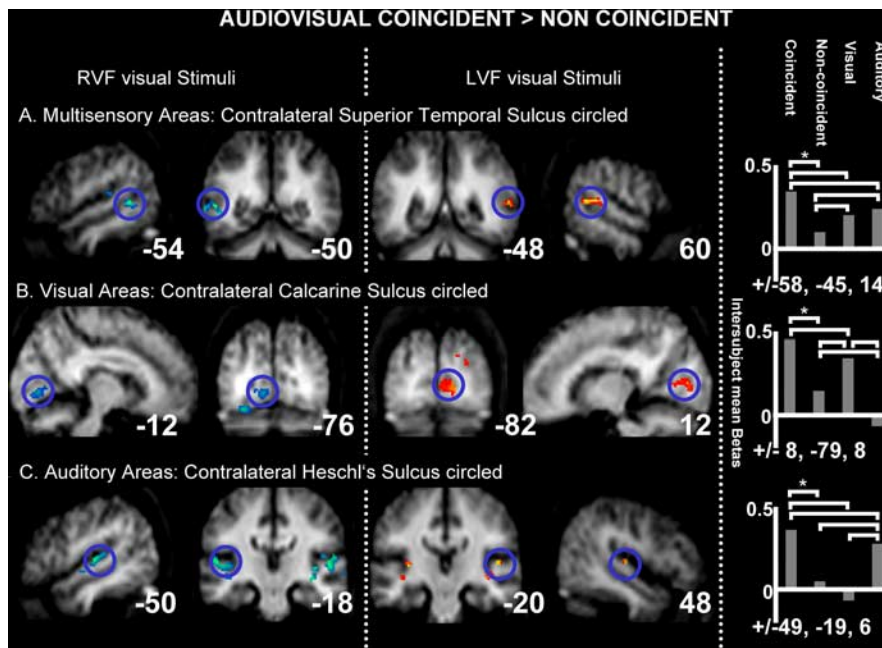
Figure 2. fMRI results: BOLD signal differences for corresponding minus noncorresponding audiovisual stimulation. Group effects in the following: **A**, contralateral multisensory superior temporal sulcus; **B**, contralateral early visual cortex; **C**, bilateral auditory cortex, with contralateral peak. Shown for the RVF and LVF groups (columns 1, 2 and 3, 4, respectively). The intersubject mean parameter estimates (SPM betas, proportional to percentage signal change) are plotted for contralateral mSTS, primary visual cortex, and primary auditory cortex (each plot in corresponding rows to the brain activations shown) from the subject-specific maxima used in the individual analyses, averaged across LVF and RVF groups, with mean Montreal Neurological Institute coordinates below each bar graph. Brackets linking pairs of bars in these graphs all indicate significant differences across those conditions ($p < 0.05$ or better).

0.001) (Fig. 2*B*, Table 2*B*). Again, this effect was found for each individual subject, in the anterior lower lip of their calcarine fissure (for illustrative single subject, see Fig. 3*B*) representing the contralateral peripheral upper visual quadrant, in which the visual stimuli appeared.

Finally, enhanced BOLD response for AVC > NC stimulation was found also within sensory-specific auditory cortex, in the vicinity of Heschl's gyrus, also peaking contralateral to the coincident visual hemifield (peaks at -48, -20, 10 for RVF and 50, -16, 8 for LVF group; both p values < 0.001) (Fig. 2*C*, Table 2*C*), albeit with some bilateral activations also found yet a systematically contralateral peak. We found this pattern in 23 of the 24 individual subjects, within the medial part of anteriormost Heschl's gyrus (typically considered as primary auditory cortex), often extending into posterior insula and planum temporale (for illustrative single subject, see Fig. 3*C*).

Mean parameter estimates (SPM betas, proportional to percentage signal change) from individual peaks in contralateral calcarine sulcus and contralateral Heschl's gyrus are plotted in Figure 2, *B* and *C* (bar graphs), respectively. In addition to the clear AVC > NC effect, AVC also elicited a higher BOLD signal than the relevant unisensory baseline (i.e., vision for calcarine sulcus, auditory for Heschl's gyrus; each at $p < 0.008$ or better), whereas the NC condition was significantly lower than those unisensory baselines ($p < 0.007$ or better).

Comparison of our two multisensory conditions with the unisensory baselines

Although our main focus was on comparing audiovisual correspondence minus noncorrespondence (AVC > NC), the plots in Figure 2 show that AVC also elicited higher activity than either unisensory baseline in mSTS, whereas NC was lower than both these baselines there. This might reflect corresponding auditory and visual events becoming "allies" in neural representation (attributable to their correspondence), whereas noncorresponding instead become "competitors," leading to the apparent suppression observed for them in mSTS. This might also hold for the A1 and V1 results, in which the most relevant unisensory baseline (audition or vision, respectively) was again significantly below the AVC condition yet significantly above the NC. Alternatively, one might argue that the level of activity for NC in A1 or V1 may correspond to the combined mean of the separate auditory and visual baselines for that particular area (NC did not differ from that mean for V1 and A1, although it did for STS). However, this would still imply that combining noncorresponding sounds and lights can reduce activity in primary sensory cortices relative to the preferred modality alone, although temporally corresponding audiovisual stimulation boosts activity in both V1 and A1.

Our finding of enhanced BOLD signal in the AVC condition, but reduced in the NC, relative to unisensory baselines is reminiscent in some respects of an interesting recent audiotactile (rather than audiovisual) study by Lakatos et al. (2007). Unlike the present whole-brain human fMRI method, they measured lamina-specific multiunit activity (MUA) invasively in macaque primary auditory cortex and calculated current-source density distributions (CSDs). Responses to combined audiotactile stimuli differed from summed unisensory tactile and auditory responses, indicating a modulatory influence of tactile stimuli on auditory. The stimulus onset asynchronies producing either response enhancement or suppression for multisensory stimulation hinted at a phase-resetting mechanism affecting neural oscillations. In particular, because of rapid somatosensory input into supragranular layers of A1 (at ~8 ms), corresponding auditory signals may arrive (~9 ms) at an optimal excitable phase (hence producing an enhanced response, potentially analogous to our result for the AVC condition) but may arrive during an opposing non-excitable phase when somatosensory inputs do not correspond (hence producing a depressed response, potentially analogous to our result for the NC condition, although note that enhanced audiotactile CSDs have been observed at various SOAs because of the oscillatory nature of the underlying mechanism). Although some analogies can be drawn to our fMRI results, the sluggish temporal resolution of fMRI (compared with MUA and CSD) precludes any links between our study and that of Lakatos et al. (2007) from being pushed too far. Electroencephalography (EEG) or magnetoencephalography (MEG) might be more suitable for studying the timing and oscillatory nature of the present effects, or our present paradigm could be applied to monkeys during invasive recordings (because the only task required is fixation monitoring). The architecture and timing of possible visual inputs into A1 might differ from those for somatosen-
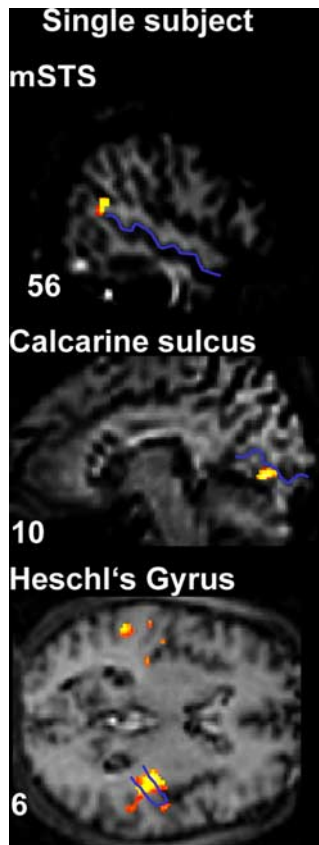
Figure 3. fMRI results: BOLD signal differences for corresponding minus noncorresponding audiovisual stimulation in an illustrative single subject. mSTS, visual cortex, and auditory cortex activations are shown, with the STS, the calcarine fissure, and Heschl's gyrus highlighted in blue on that individual's anatomical scan. Localization of the effects with respect to these anatomical landmarks was implemented in every individual.

sory (as studied by Lakatos et al., 2007) and are likely to be slower (because of retinal transduction time), probably too slow to act exactly like the somatosensory inputs in the study by Lakatos et al. Brosch et al. (2005) reported modulated MUAs in A1 starting only at ~60–100 ms after presentation of a visual stimulus. Ghazanfar et al. (2005) reported their first audiovisual interactions in A1 at ~80 ms after stimulus (see also below). Finally, here we found fMRI effects not only for A1 but also for V1 and STS. As will be seen below, analyses of effectivity suggested some possible feedback influences from STS on A1 [as had been suggested by Ghazanfar et al. (2005), their pp 5004, 5011] and on V1, for the present fMRI data in our paradigm.

Several different contrasts and analysis approaches have been introduced in previous multisensory research when comparing multisensory conditions with unisensory baselines. Although the present study focuses on fMRI measures, Stein and colleagues conducted many influential single-cell studies on the superior colliculus and other structures (Stein, 1978; Stein and Meredith, 1990; Stein et al., 1993; Wallace et al., 1993, 1996; Wallace and Stein, 1994, 1996, 1997). They suggested that, depending on the relative timing and/or location of multisensory inputs, neural responses can sometimes exceed (or fall below) the sum of the responses for each unisensory input (Lakatos et al., 2007). Nonlinear analysis criteria have also been applied to EEG data in some

multisensory studies that typically manipulated presence/absence of costimulation in a second modality (Giard and Peronnet, 1999; Foxe et al., 2000; Fort et al., 2002; Molholm et al., 2002, 2004; Murray et al., 2005) rather than a detailed relationship in temporal patterning as here. Similar nonadditive criteria have even been applied to fMRI data (Calvert et al., 2001). Conversely, such criteria have been criticized for some situations (for ERP contrasts, see Teder-Sälejärvi et al., 2002). Moreover, Stein and colleagues subsequently reported that some of the cellular phenomena that originally inspired such criteria may in fact more often reflect linear rather than nonlinear phenomena (Stein et al., 2004) when considered at the population level. Such considerations have led to proposals of revised criteria for fMRI studies of multisensory integration, including suggestions that a neural response significantly different from the maximal unisensory response may be taken to signify a multisensory effect (Beauchamp, 2005b). Most importantly, we note that the critical fMRI results reported here cannot merely reflect summing (or averaging) of two entirely separate BOLD responses to incoming auditory and visual events, because otherwise the outcome should have been comparable for corresponding and noncorresponding conditions. Recall that the auditory and visual stimuli themselves were equivalent and fully counterbalanced across our AVC and NC conditions; only their temporal relationship varied. Hence, our critical effects must reflect multisensory effects that depend on the temporal correspondence of incoming auditory and visual temporal patterns.

Analysis of functional changes in effective connectivity between brain areas

Given the activation results, we seeded our PPI connectivity analysis at mSTS (Fig. 4D, blue region) in a spherical region (4 mm diameter) surrounding the maximum found in the main analyses for each individual (for coordinates of the group average, see Fig. 4A). This PPI analysis revealed that functional coupling of seeded mSTS, contralateral to the crossmodal coincidence, was specifically enhanced (showed stronger covariation) with early visual cortex (mean peak coordinates, ±4, −82, 6; $p < 0.008$) (Fig. 4E) and auditory cortex (±44, 22, 6; $p < 0.02$) (Fig. 4F) ipsilaterally to the mSTS seed in the context of audiovisual coincidence (vs noncoincidence). This modulation is not redundant with the overall BOLD activations reported above, because it reflects condition-dependent covariation between brain regions, after mean activations by condition for each region have been discounted (see Materials and Methods) (Friston et al., 1997). Nevertheless, these connectivity results closely resembled the activation pattern in terms of the brain regions implicated (Fig. 4, compare A–C with D–F; see also Fig. 2), providing additional evidence to highlight an mSTS–A1–V1 interconnected network for the present effects of audiovisual temporal correspondence in our paradigm. The highly specific pattern of condition-dependent functional coupling with mSTS was found in visual cortex for all 24 individual subjects and in auditory cortex for 23 of 24 subjects (for a representative subject, see Fig. 5) (for every single individual, see supplemental data, available at www.jneurosci.org as supplemental material).

Although several studies also implicate a role for multisensory thalamic nuclei in cortical response profiles (Baier et al., 2006, Lakatos et al., 2007) and the thalamus is increasingly regarded as a potentially major player in multisensory phenomena (Jones, 1998; Schroeder et al., 2003; Fu et al., 2004; Ghazanfar and Schroeder, 2006), we did not observe any BOLD effects in the thala-
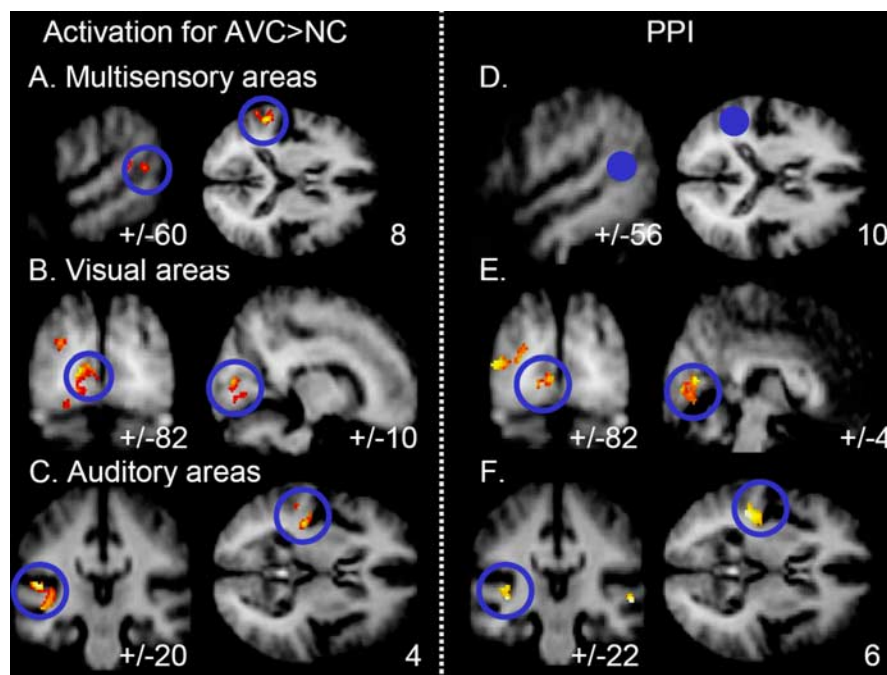
Figure 4. ***A–C***, Combined results of corresponding minus noncorresponding audiovisual stimulation for LVF and RVF groups, with hemisphere flipping to pool results contralateral to the audiovisual coincidence, which thereby appear in the apparently left hemisphere here (see Materials and Methods). Overall activations for AVC > NC in the following: ***A***, contralateral mSTS; ***B***, contralateral early visual cortex; ***C***, contralateral auditory cortex. ***D–F***, Enhanced functional coupling of seeded mSTS (this seeded region shown in ***D*** as filled blue circle) with visual (***E***) and auditory (***F***) areas in the context of audiovisual temporal coincidence versus noncoincidence. Voxels showing significantly greater functional coupling with the STS seed for that context are highlighted in red.

mus with the human fMRI method used here, only cortically. Although fMRI may not be an ideal method for detecting any effects in the thalamus (particularly if subtle or layer specific), this does not undermine our positive cortical findings. Moreover, new fMRI methods are now being developed to enhance sensitivity to subcortical thalamic structures (Schneider et al., 2004) and, as noted previously, our new paradigm may also be suitable for invasive animal work in the future because the only task required is fixation monitoring.

Because PPI analyses are nondirectional in nature (see Materials and Methods), we further assessed possible influences between mSTS, V1, and A1 for the present fMRI data using a DIT measure (Hinrichs et al., 2006). Data from A1 and V1 were derived from the subject-specific maxima for overlap between the basic activation analysis and the PPI analysis (mean coordinates: V1, ±8.9, −78.4, 6.9; A1, ±47.6, 19.7, 7.1). Inferred information flow from mSTS toward V1 and toward A1 was significantly higher than the opposite direction during audiovisual temporal coincidence ($p < 0.05$ in both cases) (Fig. 6) relative to temporal noncoincidence. No reliable condition-specific differences were found for any direct A1–V1 influences.

Thus, visual and auditory cortices not only showed activation by audiovisual temporal correspondence in the present fMRI data; over and above this, they also showed some functional coupling with mSTS, as confirmed when seeding the PPI analysis there revealed condition-specific effective connectivity with A1 and V1. Moreover, DIT analysis suggested a significantly increased influence from mSTS on A1 and V1 specifically during audiovisual temporal correspondence, rather than direct A1–V1 influences, for these fMRI data.

As noted previously, possible thalamic influences on mul-

tisensory effects (Lakatos et al., 2007) may also need to be considered and may not be readily detected with human fMRI, although this does not undermine the positive evidence we did find for feedback influences from mSTS on A1 [as also hypothesized by Ghazanfar et al. (2005), their pp 5004, 5011]. More generally, the relative balance between bottom-up multisensory influences (e.g., via the thalamus, or cortical–cortical as between A1 and V1) and top-down feedback influences (as suggested here, by the DIT analysis, for STS influences on A1 and V1 in the AVC condition) may depend on the paradigm used. It is possible that, in our paradigm, temporal correspondence between auditory and visual streams tended to attract some attention to those streams, which might favor feedback influences such as the DIT effect we observed from STS on A1 and V1. Conversely, the AVC condition did not activate those brain structures (including parietal cortex and more anterior regions) that are classically associated with attention capture (Corbetta and Shulman, 2002; Watkins et al., 2007). Moreover, performance of the central task (which was off ceiling at ~83% correct) did not suffer in the AVC condition, whereas an attention-capture account might have predicted a decrement for that. Nevertheless, we return to consider the possible attention issue below.

Discussion
We found with human fMRI that AVC in temporal pattern can affect not only brain regions traditionally considered to be multisensory, as for contralateral mSTS, but also sensory-specific visual and auditory cortex, including even primary cortices. This impact of AVC was systematically contralateral to the peripheral stimuli, ruling out nonspecific explanations such as higher arousal in one condition than another. Contralateral preferences for STS accord with some animal single-cell work (Barraclough et al., 2005).

Activation of contralateral human mSTS by audiovisual temporal correspondence
A role for STS in audiovisual integration would accord generally with single-cell studies (Benevento et al., 1977; Bruce et al., 1981; Barraclough et al., 2005), lesion data (Petrides and Iversen, 1978), and other human neuroimaging work (Miller and D'Esposito, 2005; van Atteveldt et al., 2006; Watkins et al., 2006) that typically used more complex or semantic stimuli than here. However, to our knowledge, no previous human study has observed the systematic contralaterality found here, nor the clear effects on primary visual and auditory cortex in addition to mSTS, attributable solely to temporal correspondence between simple flashes and beeps (although for potentially related monkey A1 studies, see Ghazanfar et al., 2005; Lakatos et al., 2007), nor the informative pattern of functional coupling that we observed.
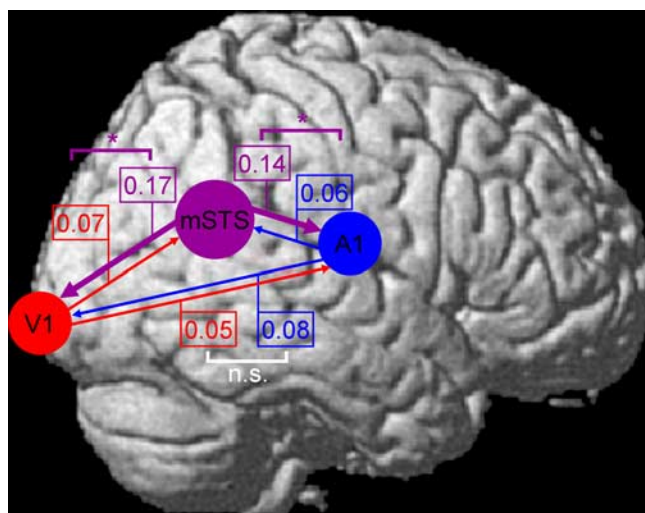
Calvert et al. (2001) implicated human STS in audiovisual
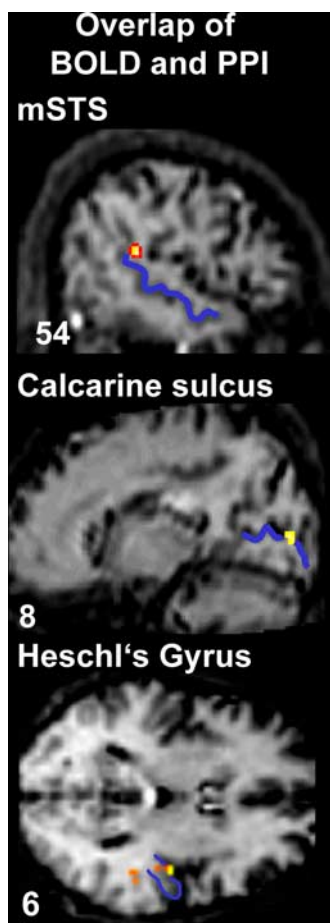
Figure 5. Overlap of observed activation, for audiovisual correspondence minus noncorrespondence with functional coupling results from one illustrative participant, with STS, calcarine fissure, and Heschl's gyrus draw in blue onto the individual's structural scan. STS activation was used as the seed for the PPI analysis, whereas regions in Heschl's gyrus and calcarine fissure show both increased activation for audiovisual correspondence minus noncorrespondence and also enhanced coupling with ipsilateral STS in the context of audiovisual temporal correspondence [this overlap was formally tested for each individual by a conjunction of PPI results and experimental AVC minus NC within areas that showed a sensory-specific effect (visual > auditory and auditory > visual, respectively)].

Figure 6. Results of directed information transfer analysis for temporally corresponding minus noncorresponding audiovisual conditions (i.e., difference in the inferred directional information transfer, attributable to condition) between STS (indicated schematically with purple circle), calcarine fissure (V1, indicated schematically with red circle), and Heschl's gyrus (A1, schematic blue circle), with direction of information transfer indicated via colored arrows. Numbers by each arrow indicate the measured change in influence (larger = stronger for temporally corresponding than noncorresponding condition) in the direction of each colored arrow. Colored brackets link pairs of numbers showing significant differences between the impact of condition, indicating that one direction of influence changed more than the converse direction attributable to temporal correspondence ($p < 0.05$ or better). White brackets indicate no such significant differences [nonsignificant (n.s.)]. The absolute values for DIT measures matter less than the reliability of any differences because absolute values can depend on imaging parameters (Hinrichs et al., 2006).

integration via neuroimaging, when using analysis criteria derived from classic electrophysiological work. Several human fMRI studies used other criteria to relate STS to audiovisual integration, for semantically related objects and sounds (Beauchamp et al., 2004a; Miller and D'Esposito, 2005; van Atteveldt et al., 2006). However, here we manipulated only temporal correspondence between meaningless flashes and beeps but ensured that all other temporal factors were held constant (unlike studies that compared, say, rhythmic with arrhythmic stimuli). Atteveldt et al. (2006) varied temporal alignment plus semantic congruency between visual letter symbols and auditory phonemes, reporting effects in anterior STS, but their paradigm did not assess crossmodal relationships in rapid temporal patterning (i.e., their letters did not correspond in temporal structure to their speech sounds). Several previous imaging studies (Bushara et al., 2001; Bischoff et al., 2007; Dhamala et al., 2007) used tasks explicitly requiring subjects to judge temporal audiovisual (a)synchrony versus some other task but may thereby have activated task-related networks rather than stimulus-driven modulation as here. Although our results converge with a wide literature in implicating human

STS in audiovisual interactions, they go beyond this in showing specifically contralateral activations, determined solely by audiovisual temporal correspondence for nonsemantic stimuli, while identifying interregional functional coupling.

Several previous single-cell studies considered the temporal window of multisensory integration for a range of brain areas (Meredith et al., 1987; Avillac et al., 2005; Lakatos et al., 2007). Here the average stimulus rate was 4 Hz; although this might be extracted by the brain, the streams were in fact highly arrhythmic, with a rectangular distribution of 2–10 Hz and no particularly prominent temporal frequency when Fourier transformed. The minimally protective temporal window separating auditory and visual events when noncorresponding was 100 ms. Such temporal constraints were evidently sufficient to modulate human mSTS (plus A1 and V1; see below) in a highly systematic manner.

One interesting question for the future is whether the present effects of AVC may evolve and increase over the course of an ongoing stream, as might be expected if they reflect some entrainment (Lakatos et al., 2005) of neural oscillations, possibly involving a reset of delta/theta frequency-band modulations (Lakatos et al., 2007). fMRI as used here may be less suitable than EEG/MEG, or invasive recordings in animals, for resolving this. Presumably, any such entrainment mechanisms might be more pronounced for rhythmic stimulus trains (Lakatos et al., 2005) than with highly arrhythmic streams as here. Conversely, these erratic, arrhythmic streams may provide particularly strong information that auditory and visual events are related when they perfectly correspond, because this is highly unlikely to arise by chance for such irregular events.

Effects of audiovisual temporal correspondence on primary sensory areas and on functional connectivity between areas

In addition to mSTS, we found that sensory-specific visual and auditory cortex (including V1 and A1) showed effects of audiovisual temporal correspondence, primarily contralateral to the visual stream. This pattern was confirmed in all 24 individuals (except one for A1), indicating that multisensory factors can affect human brain regions traditionally considered unisensory. This has become an emerging theme in recent multisensory work, using different neural measures (cf. Giard and Peronnet, 1999; Macaluso et al., 2000; Molholm et al., 2002; Brosch et al., 2005; Ghazanfar et al., 2005; Miller and D'Esposito, 2005; Watkins et al., 2006; Kayser et al., 2007; Lakatos et al., 2007).

Several aspects of neuroanatomical architecture have been considered as potentially contributing to multisensory interplay (Schroeder and Foxe, 2002; Schroeder et al., 2003), including feedforward thalamocortical, direct cortical–cortical links between modality-specific areas, or feedback influences. The V1 and A1 effects observed here might reflect backprojections from mSTS, for which there is anatomical evidence in animals (Falchier et al., 2002). Alternatively, they might in principle reflect direct V1–A1 connections or thalamic modulation (although we found no significant thalamic effects here, possibly attributable to limits of fMRI). Some evidence for A1–V1 connections has been found in animal anatomy, although these appear sparse compared with connections involving mSTS (Falchier et al., 2002). Some human ERP evidence for early multisensory interactions involving auditory (and tactile) stimuli, which may arise in sensory-specific cortices, has been reported (Murray et al., 2005), as have some fMRI-modulations in high-resolution monkey studies (Kayser et al., 2007), and differential MUAs/CSDs in monkey A1 (Lakatos et al., 2007).

Here we approached the issue of inter-regional influences with human fMRI data using two established analysis approaches to functional coupling or "connectivity" between regions: the PPI approach and the DIT approach. PPI analysis revealed significantly enhanced coupling of seeded mSTS with ipsilateral V1 and A1, specific to the AVC condition. The DIT analysis revealed significantly higher "information flow" from mSTS to both A1 and V1 than in the opposite direction during the AVC condition relative to the NC condition. DIT measures for "direct" influences between A1 and V1 found no significant impact of audiovisual temporal correspondence versus noncorrespondence. This appears consistent with mSTS modulating A1 and V1 when auditory and visual inputs correspond temporally. This issue could be addressed further with neural measures that have better temporal resolution (e.g., EEG/MEG, or invasive animal recordings in a similar paradigm). It should also be considered whether possible "attention capture" by corresponding streams could contribute to feedback influences predominating, as mentioned previously. Any audiovisual temporal correspondence was always task irrelevant here, performance of the central task did not vary with peripheral experimental condition, and brain regions conventionally associated with attention shifts were not activated by AVC. Nevertheless, increasing attentional load for the central task (Lavie, 2005) might conceivably modulate the present effects.

The hypothesis of feedback influences from mSTS, to A1 in particular, was suggested by Ghazanfar et al. (2005), who reported increased neural responses within monkey A1 for audiovisually congruent (and thus temporally correspondingly although also semantically matching) monkey vocalizations. Those authors hypothesized (their pp 5004, 5011) that A1 enhancement might reflect feedback from STS, as also suggested by the very different type of evidence here. Animal work suggests that visual input into auditory belt areas arrives at the supragranular layer, in apparent accord with a feedback loop, although other neighboring regions in and around auditory cortex evidently do receive direct somatosensory afferents plus inputs from multisensory thalamic nuclei (Schroeder et al., 2003; Ghazanfar and Schroeder, 2006; Lakatos et al., 2007).

For the present human paradigm, the idea of feedback from mSTS on visual and auditory cortex might be tested directly by combining our fMRI paradigm with selective lesion/transcranial magnetic stimulation work. If mSTS imposes the effects on A1 and V1, a lesion in mSTS should presumably eliminate these effects within intact ipsilateral A1 and V1. In contrast, if direct A1–V1 connections or thalamocortical circuits are involved, effects of audiovisual temporal correspondence on V1/A1 should remain unchanged. Finally, because our new paradigm uses simple nonsemantic stimuli (flashes and beeps) and only requires a fixation-monitoring task, it could be applied to nonhuman primates to enable more invasive measures to identify the pathways and mechanisms. A recent monkey study on audiovisual integration (Kayser et al., 2007) introduced promising imaging methods for such an approach, whereas Ghazanfar et al. (2005) and Lakatos et al. (2007) illustrate the power of invasive recordings.

Conclusion

Our fMRI results show that audiovisual correspondence in temporal patterning modulates contralateral mSTS, A1, and V1. This confirms in humans that multisensory relationships can affect not only conventional multisensory brain structures (as for STS) but also primary sensory cortices when auditory and visual inputs have a related temporal structure that is very unlikely to arise by chance alone and is therefore highly likely to reflect a common source in the external world.

References

Avillac M, Deneve S, Olivier E, Pouget A, Duhamel JR (2005) Reference frames for representing visual and tactile locations in parietal cortex. Nat Neurosci 8:941–949.

Baier B, Kleinschmidt A, Müller NG (2006) Cross-modal processing in early visual and auditory cortices depends on expected statistical relationship of multisensory information. J Neurosci 26:12260–12265.

Barraclough NE, Xiao DK, Baker CI, Oram MW, Perrett DI (2005) Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. J Cogn Neurosci 17:377–391.

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

Beauchamp MS (2005b) Statistical criteria in FMRI studies of multisensory integration. Neuroinformatics 3:93–113.

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

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

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.

Bischoff M, Walter B, Blecker CR, Morgen K, Vaitl D, Sammer G (2007) Utilizing the ventriloquism-effect to investigate audio-visual binding. Neuropsychologia 45:578–586.

Brosch M, Scheich H (2005) Non-acoustic influence on neural activity in auditory cortex. In: Auditory cortex: a synthesis of human and animal

research (König P, Heil P, Budinger E, Scheich H, eds), pp 127–143. Hillsdale, NJ: Erlbaum.

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, Bullmore ET, Brammer MJ, Campbell R, Williams SC, McGuire PK, Woodruff PW, Iversen SD, David AS (1997) Activation of auditory cortex during silent lipreading. Science 276:593–596.

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.

Calvert GA, Stein BE, Spence C (2004) The handbook of multisensory processing. Cambridge, MA: MIT.

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

Cusick CG (1997) The superior temporal polysensory region in monkeys. In: Cerebral cortex: extrastriate cortex in primates (Rockland KS, Kaas JH, Peters A, eds), pp 435–463. New York: Plenum.

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

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

Fort A, Delpuech C, Pernier J, Giard MH (2002) Early auditory-visual interactions in human cortex during nonredundant target identification. Brain Res Cogn Brain Res 14:20–30.

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

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. Brain Res Cogn Brain Res 10:77–83.

Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ (1997) Psychophysiological and modulatory interactions in neuroimaging. NeuroImage 6:218–229.

Fu KM, Shah AS, O'Connell MN, McGinnis T, Eckholdt H, Lakatos P, Smiley J, Schroeder CE (2004) Timing and laminar profile of eye-position effects on auditory responses in primate auditory cortex. J Neurophysiol 92:3522–3531.

Fujisaki W, Koene A, Arnold D, Johnston A, Nishida S (2006) Visual search for a target changing in synchrony with an auditory signal. Proc Biol Sci 273:865–874.

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.

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

Hinrichs H, Heinze HJ, Schoenfeld MA (2006) Causal visual interactions as revealed by an information theoretic measure and fMRI. NeuroImage 31:1051–1060.

Jones EG (1998) Viewpoint: the core and matrix of thalamic organization. Neuroscience 85:331.

Kaas JH, Collins CE (2004) The resurrection of multisensory cortex in primates. In: The handbook of multisensory processes (Calvert GA, Spence C, Stein BE, eds), pp 285–294. Cambridge, MA: Bradford.

Kanowski M, Rieger JW, Noesselt T, Tempelmann C, Hinrichs H (2007) Endoscopic eye tracking system for fMRI. J Neurosci Methods 160:10–15.

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.

Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE (2005) An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J Neurophysiol 94:1904–1911.

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

Lavie N (2005) Distracted and confused?: selective attention under load. Trends Cogn Sci 9:75–82.

Lipschutz B, Kolinsky R, Damhaut P, Wikler D, Goldman S (2002) Attention-dependent changes of activation and connectivity in dichotic listening. NeuroImage 17:643–656.

Macaluso E, Driver J (2005) Multisensory spatial interactions: a window onto functional integration in the human brain. Trends Neurosci 28:264–271.

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

Macaluso E, George N, Dolan R, Spence C, Driver J (2004) Spatial and temporal factors during processing of audiovisual speech: a PET study. NeuroImage 21:725–732.

McDonald JJ, Teder-Salejarvi WA, Hillyard SA (2000) Involuntary orienting to sound improves visual perception. Nature 407:906–908.

McDonald JJ, Teder-Salejarvi WA, Di Russo F, Hillyard SA (2003) Neural substrates of perceptual enhancement by cross-modal spatial attention. J Cogn Neurosci 15:10–19.

Meredith MA, Nemitz JW, Stein BE (1987) Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. J Neurosci 7:3215–3229.

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

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. Brain Res Cogn Brain Res 14:115–128.

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

Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, Schroeder CE, Foxe JJ (2005) Grabbing your ear: rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. Cereb Cortex 15:963–974.

Noesselt T, Fendrich R, Bonath B, Tyll S, Heinze HJ (2005) Closer in time when farther in space—spatial factors in audiovisual temporal integration. Brain Res Cogn Brain Res 25:443–458.

Petrides M, Iversen SD (1978) The effect of selective anterior and posterior association cortex lesions in the monkey on performance of a visual-auditory compound discrimination test. Neuropsychologia 16:527–537.

Schneider KA, Richter MC, Kastner S (2004) Retinotopic organization and functional subdivisions of the human lateral geniculate nucleus: a high-resolution functional magnetic resonance imaging study. J Neurosci 24:8975–8985.

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.

Schroeder CE, Smiley J, Fu KG, McGinnis T, O'Connell MN, Hackett TA (2003) Anatomical mechanisms and functional implications of multisensory convergence in early cortical processing. Int J Psychophysiol 50:5–17.

Slutsky DA, Recanzone GH (2001) Temporal and spatial dependency of the ventriloquism effect. NeuroReport 12:7–10.

Spence C, Driver J (2004) Crossmodal space and crossmodal attention. New York: Oxford UP.

Stein BE (1978) Development and organization of multimodal representation in cat superior colliculus. Fed Proc 37:2240–2245.

Stein BE, Meredith MA (1990) Multisensory integration. Neural and behavioral solutions for dealing with stimuli from different sensory modalities. Ann NY Acad Sci 608:51–65, discussion 65–70.

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

Stein BE, Meredith MA, Wallace MT (1993) The visually responsive neuron and beyond: multisensory integration in cat and monkey. Prog Brain Res 95:79–90.

Stein BE, Stanford TR, Wallace MT, Vaughan JW, Jiang W (2004) Cross-modal spatial interactions in subcortical and cortical circuits. In: Cross-

modal space and crossmodal attention (Spence C, Driver J, eds), pp 25–50. Oxford: Oxford UP.

Teder-Sälejärvi WA, McDonald JJ, Di Russo F, Hillyard SA (2002) An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. Cogn Brain Res 14:106–114.

van Atteveldt NM, Formisano E, Blomert L, Goebel R (2006) The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. Cereb Cortex 17:962–974.

Wallace MT, Stein BE (1994) Cross-modal synthesis in the midbrain depends on input from cortex. J Neurophysiol 71:429–432.

Wallace MT, Stein BE (1996) Sensory organization of the superior colliculus in cat and monkey. Prog Brain Res 112:301–311.

Wallace MT, Stein BE (1997) Development of multisensory neurons and multisensory integration in cat superior colliculus. J Neurosci 17:2429–2444.

Wallace MT, Meredith MA, Stein BE (1993) Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. J Neurophysiol 69:1797–1809.

Wallace MT, Wilkinson LK, Stein BE (1996) Representation and integration of multiple sensory inputs in primate superior colliculus. J Neurophysiol 76:1246–1266.

Watkins S, Shams L, Tanaka S, Haynes JD, Rees G (2006) Sound alters activity in human V1 in association with illusory visual perception. NeuroImage 31:1247–1256.

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