Short-Term Memory for Space and Time Flexibly Recruit Complementary Sensory-Biased Frontal Lobe Attention Networks

Highlights

- fMRI reveals interleaved auditory and visual attention networks in human frontal lobe

- Modality-specific functional connectivity links frontal and posterior cortex

- Spatial short-term memory recruits frontal vision areas, even for auditory stimuli

- Timing short-term memory recruits frontal auditory areas, even for visual stimuli

Authors

Samantha W. Michalka, Lingqiang Kong, Maya L. Rosen, Barbara G. Shinn-Cunningham, David C. Somers

Correspondence

samantha.michalka@gmail.com (S.W.M.),
somers@bu.edu (D.C.S.)

In Brief

Michalka et al. report four interleaved vision-biased and auditory-biased attention regions bilaterally in human lateral frontal cortex. Short-term memory for space and for time recruits the frontal visual and auditory networks, respectively across sensory modalities.

CrossMark

CellPress

Short-Term Memory for Space and Time Flexibly Recruit Complementary Sensory-Biased Frontal Lobe Attention Networks

Samantha W. Michalka,[1,3,*] Lingqiang Kong,[1,2] Maya L. Rosen,[2] Barbara G. Shinn-Cunningham,[1,4] and David C. Somers[1,2,3,*]
[1]Center for Computational Neuroscience and Neural Technology, Boston University, Boston, MA 02215, USA
[2]Department of Psychological and Brain Sciences, Boston University, Boston, MA 02215, USA
[3]Graduate Program for Neuroscience, Boston University, Boston, MA 02215, USA
[4]Biomedical Engineering, Boston University, Boston, MA 02215, USA
*Correspondence: samantha.michalka@gmail.com (S.W.M.), somers@bu.edu (D.C.S.)
http://dx.doi.org/10.1016/j.neuron.2015.07.028

SUMMARY

The frontal lobes control wide-ranging cognitive functions; however, functional subdivisions of human frontal cortex are only coarsely mapped. Here, functional magnetic resonance imaging reveals two distinct visual-biased attention regions in lateral frontal cortex, superior precentral sulcus (sPCS) and inferior precentral sulcus (iPCS), anatomically interdigitated with two auditory-biased attention regions, transverse gyrus intersecting precentral sulcus (tgPCS) and caudal inferior frontal sulcus (cIFS). Intrinsic functional connectivity analysis demonstrates that sPCS and iPCS fall within a broad visual-attention network, while tgPCS and cIFS fall within a broad auditory-attention network. Interestingly, we observe that spatial and temporal short-term memory (STM), respectively, recruit visual and auditory attention networks in the frontal lobe, independent of sensory modality. These findings not only demonstrate that both sensory modality and information domain influence frontal lobe functional organization, they also demonstrate that spatial processing co-localizes with visual processing and that temporal processing co-localizes with auditory processing in lateral frontal cortex.
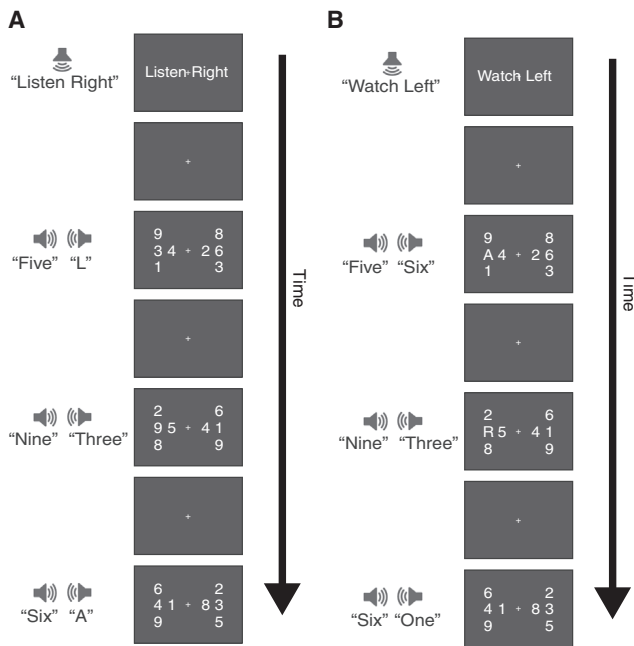
INTRODUCTION

The visual and auditory systems are each capable of coding spatial information and timing information, but they exhibit complementary strengths and weaknesses. The visual system excels at encoding spatial information—the retina records spatial information with high precision, and over 20 cortical areas exhibit visuospatial maps (Swisher et al., 2007; Wandell et al., 2007; Silver and Kastner, 2009); however, the timing of visual responses is sluggish and is influenced by nontemporal stimulus properties such as contrast (Gawne, 2000). Conversely, the auditory system codes temporal information with high resolution and utilizes very precise spike timing information, particularly in early, subcortical portions of the auditory pathway (e.g., Joris et al., 1994; Agmon-Snir et al., 1998; Adams, 2006); however, spatial information is not encoded at the cochlea; rather, it must be computed at a higher stage, and no evidence for auditory spatial maps within the cortex has been reported. These complementary strengths and weaknesses are well known within the perceptual literature. The modality appropriateness hypothesis suggests that each sensory modality is capable of a variety of functions, but is better than other modalities at certain functions; when sensory modalities conflict, the modality most "appropriate" or reliable for the particular function will dominate (Welch and Warren, 1980; O'Connor and Hermelin, 1972; Alais and Burr, 2004). Typically, when visual and auditory inputs compete, visual cues are weighted more heavily in spatial perception (Pick et al., 1969), while auditory cues are weighted more than visual cues in temporal perception (for example, see Welch et al., 1986; Shams et al., 2000; Recanzone, 2003). Behavioral evidence suggests that unisensory short-term memory can leverage these specializations; specifically, unisensory inputs may be cross-modally encoded into the short-term memory representations associated with the "appropriate" modality (e.g., "hearing visual rhythms" in Guttman et al., 2005).

We hypothesize that (1) higher-order cortical structures exhibit strong biases for attention to either visual or auditory information, (2) these structures functionally link information domain (time or space) with the "appropriate" sensory modality (spatial/vision; temporal/audition), and (3) sensory information from the "inappropriate" modality can flexibly recruit these structures when a task demands high functioning in the nonpreferred information domain (i.e., spatial or temporal). We call this neural hypothesis the domain recruitment hypothesis. Here, we performed a series of fMRI experiments to test the components of the domain recruitment hypothesis and to investigate visual and auditory processing in human lateral frontal cortex.

Sensory modality is a primary organizing feature of posterior cortical regions; however, the role of sensory modality in frontal lobe organization remains controversial. While one recent multivariate analysis indicated that posterior lateral frontal cortex contains information reflecting input sensory modality

A

"Listen Right" | Listen-Right

"Five" "L"

"Nine" "Three"

"Six" "A"

B

"Watch Left" | Watch Left

"Five" "Six"

"Nine" "Three"

"Six" "One"

Figure 1. Experiment 1 Paradigm

Task schematic for visual versus auditory sustained spatial attention task (Experiment 1) showing examples of (A) attend auditory and (B) attend visual conditions. Each block began with an instruction for the subject to attend to one of four serial presentation streams (listen/watch, left/right). Subjects monitored the cued stream and reported the identity of digits (1–4) while ignoring distracting letters (in attended stream) and digits (in all other streams). Visual streams included six additional distractor streams to balance task difficulty between auditory and visual streams.

(Tamber-Rosenau et al., 2013), prior human univariate functional magnetic resonance imaging (fMRI) studies of vision and audition either point to shared multisensory structures in lateral frontal cortex (Lewis et al., 2000; Johnson and Zatorre, 2006; Ivanoff et al., 2009; Karabanov et al., 2009; Tark and Curtis, 2009; Tombu et al., 2011; Braga et al., 2013) or report a lateral frontal cortical bias for only one modality (for example, see Crottaz-Herbette et al., 2004; Jantzen et al., 2005; Rämä and Courtney, 2005; Salmi et al., 2007), which could reflect differences in task difficulty rather than sensory modality. Studies in nonhuman primates have reported distinct areas in lateral frontal cortex that are biased toward audition or vision in anatomical connectivity and/or functional response (for example, see Barbas and Mesulam, 1981; Petrides and Pandya, 1999; Romanski and Goldman-Rakic, 2002; Romanski, 2007).

Our first two experiments investigate whether sensory modality is a determining factor in the functional organization of lateral frontal cortex. The first experiment manipulates attention to sensory modality and reveals two visual-biased regions interleaved with two auditory-biased regions in lateral frontal cortex. The second experiment confirms the observation of interleaved visual-biased and auditory-biased attention networks in lateral frontal cortex using resting-state functional connectivity. Our final two experiments investigate the domain recruitment hypothesis. In order to demonstrate flexible recruitment, the exper-

iments focus on information in a single sensory modality at a time, contrasting high spatial and high temporal demands first within purely visual tasks and then within purely auditory tasks. The results of these experiments support the domain recruitment hypothesis, revealing strong recruitment of the auditory-biased frontal regions by the visual temporal task and strong recruitment of the visual-biased frontal areas by the auditory spatial task.
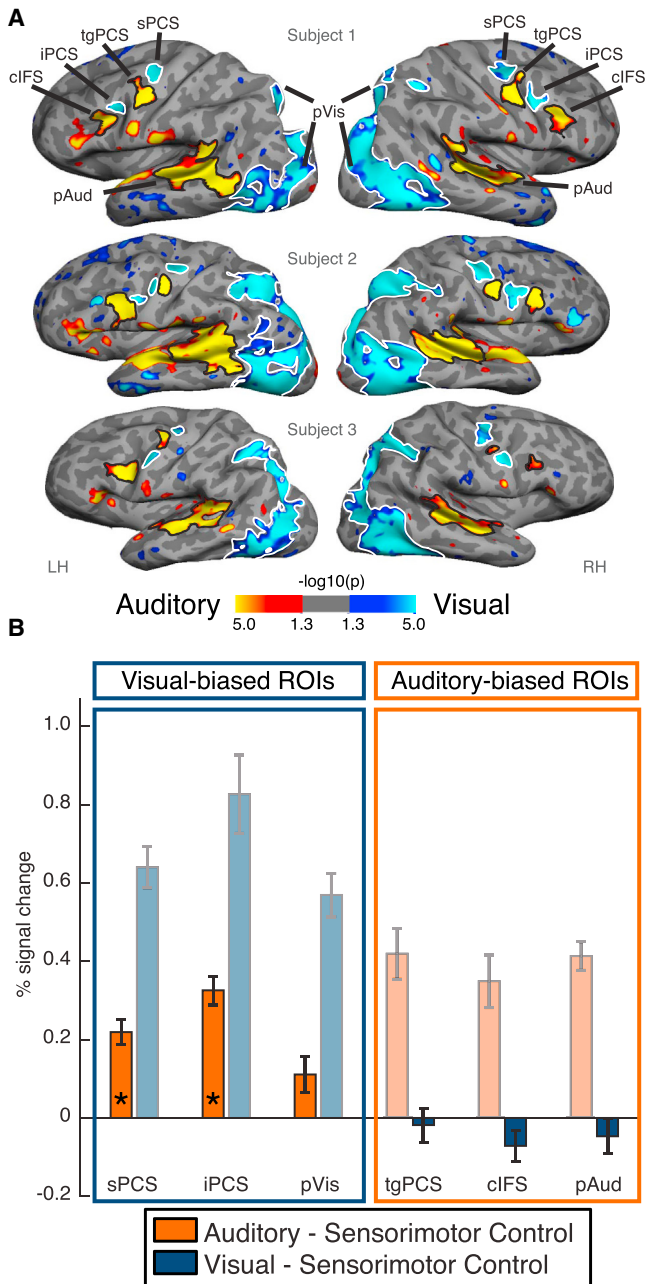
RESULTS

We performed four fMRI experiments: (1) direct comparison of sustained visual and auditory spatial attention, (2) resting-state functional connectivity using regions of interest (ROIs) defined from Experiment 1, (3) two attentionally demanding visual short-term memory tasks differing in their spatial and temporal demands, and (4) two attentionally demanding auditory short-term memory tasks differing in their spatial and temporal demands. Together, Experiments 3 and 4 served as a two-by-two investigation to dissociate processing specific to sensory modality (visual/auditory) from that specific to information domain (spatial/temporal). Eleven participants completed all four experiments; however, one participant was excluded from analysis due to excessive head movements.

Experiment 1: Sustained Visual and Auditory Spatial Attention

Participants were instructed to monitor one of four informational streams (visual left, visual right, auditory left, auditory right) and press a button when they detected a digit (a rare event among letters) in that stream while ignoring digits presented at all times in the competing streams (see Figure 1). Subjects performed at 84.1% ± 12.7% correct for visual attention blocks, and 79.9% ± 12.9% correct for auditory attention blocks with no significant difference in task performance ($t_9 = 0.94$, $p = 0.37$), indicating they successfully monitored the correct stream in both conditions.

In the caudal lateral frontal cortex of each hemisphere, a direct contrast of fMRI activation across the attended sensory modalities revealed two regions strongly biased for visual attention, interleaved with two regions strongly biased for auditory attention (see Figure 2A, Table 1, and Figure S1 available online). The superior precentral sulcus (sPCS) and inferior precentral sulcus (iPCS) exhibited a stronger blood-oxygen-level dependent (BOLD) response for visual compared to auditory sustained attention. This contrast identified the left sPCS in eight of ten subjects, the right sPCS in eight of ten subjects, and the iPCS in both the left and right hemispheres of nine of ten subjects. We consistently observed a gap between these two visual-biased areas; within this gap we observed a significant bias for sustained attention to auditory over visual stimuli. In humans, the precentral sulcus divides into two or more sections (Ono et al., 1990). The gap we observed was located where the precentral sulcus is divided by a transverse gyrus connecting the middle frontal gyrus and precentral gyrus; we henceforth refer to this area as the transverse gyrus dividing the precentral sulcus (tgPCS). The fMRI contrast of auditory greater than visual attention identified the tgPCS in the left and right hemispheres of all

A

Subject 1

tgPCS sPCS
iPCS
cIFS
pVis
pAud

sPCS tgPCS
iPCS
cIFS
pAud

Subject 2

Subject 3

LH   RH

-log10(p)

Auditory   Visual
5.0   1.3   1.3   5.0

B

Visual-biased ROIs   Auditory-biased ROIs

% signal change

1.0

0.8

0.6

0.4

0.2

0

-0.2

sPCS   iPCS   pVis   tgPCS   cIFS   pAud

■ Auditory - Sensorimotor Control
■ Visual - Sensorimotor Control

Figure 2. Contrast of Visual and Auditory Sustained Spatial Attention
(A) Statistical maps of three individual subjects showing significant differences in a direct contrast of blocks of auditory (hot colors) versus visual (cool colors) sustained spatial attention are overlaid onto cortical surface renderings (sulcus, dark gray; gyrus, light gray). Black and white outlines represent ROI definitions for auditory- and visual-biased ROIs, respectively. Note the interdigitated pattern of auditory and visual biases in the caudal lateral frontal cortex. LH, left hemisphere; RH, right hemisphere; sPCS, superior precentral sulcus; tgPCS, transverse gyrus intersecting the precentral sulcus; iPCS, inferior precentral sulcus; cIFS, caudal inferior frontal sulcus.
(B) Average percent signal change (n = 9) relative to sensorimotor control for auditory spatial attention and visual spatial attention conditions. Statistical comparisons between auditory spatial attention and visual spatial attention conditions are not included, as ROIs are defined by their direct contrast. Error bars reflect SEM.

ten subjects. In addition to the tgPCS, we observed a more ante-roventral region—the caudal portion of the inferior frontal sulcus (cIFS)—that showed BOLD responses biased toward auditory attention. cIFS was identified by the fMRI contrast in the left and right hemispheres of nine of ten subjects. Although prior fMRI studies have reported either auditory or visual activation in caudal lateral frontal cortex, this is the first report of four interdigitated regions exhibiting alternating visual and auditory biases.

After defining ROIs based on the direct contrast of auditory spatial attention blocks versus visual spatial attention blocks, we calculated the activity in each ROI separately for auditory spatial attention and visual spatial attention compared to a sensorimotor control (see Figure 2B). This analysis included two additional posterior regions: a posterior visual attention region (pVis), including the intraparietal sulcus, transverse occipital sulcus, and ventral temporal lobe, and a posterior auditory attention region (pAud) including the superior temporal gyrus and sulcus. We excluded one subject who participated in Experiment 1 for whom we failed to observe the visual-biased ROIs. Three subjects had 1–2 hemispheric ROIs that could not be identified in Experiment 1 (total was 4 out of all 72 subject hemispheric ROIs; 9 subjects × 2 hemispheres × 4 frontal ROIs per subject hemisphere); we defined those "missing" ROIs using an event-related sustained attention task based on the same stimulus set (see Experimental Procedures for details). The visual-biased ROIs, defined by greater activity during visual than auditory spatial attention, showed significant activity for auditory spatial attention relative to sensorimotor control in two frontal ROIs (sPCS: $t_8 = 6.90$, $p = 0.0006$; iPCS: $t_8 = 8.94$, $p = 0.0001$; pVis: $t_8 = 1.87$, $p = 0.39$; Holm-Bonferonni corrected). The auditory-biased ROIs, which are defined by greater activity during auditory compared to visual spatial attention, showed no significant activity during the visual spatial attention blocks relative to a sensorimotor control (tgPCS: $t_8 = -0.42$, $p = 0.68$; cIFS: $t_8 = -1.79$, $p = 0.33$; pAud: $t_8 = -1.04$, $p = 0.65$; Holm-Bonferonni corrected). Using a fixation baseline did not qualitatively change our results, and time courses indicated a consistent activity pattern throughout the blocks (see Figures S2A and S2B). Additionally, when participants attended to contralateral stimuli, we observed bilateral contralateral bias in posterior (nonfrontal) cortex in the visual attention conditions, but not in the auditory attention conditions (see Figure S2C).

Experiment 2: Intrinsic Functional Connectivity
The interdigitated pattern of visual- and auditory-biased attention regions in the caudal lateral frontal cortex found in Experiment 1 suggests that these frontal regions may be part of two distinct attention networks. To investigate the network specificity of sPCS, iPCS, tgPCS, and cIFS for visual and auditory attention, we examined their intrinsic (resting-state) functional connectivity with two posterior cortical areas, pVis and pAud. Using these seeds, defined from data in Experiment 1 (same ROIs as used for Figure 2B), we calculated seed-to-seed functional connectivity for separate resting-state fMRI runs collected in Experiment 2.

The results revealed remarkably specific intrinsic functional connectivity (see Figure 3 and Table 1). In both hemispheres

Table 1. Description of Regions of Interest Defined in Experiment 1 and Resting-State Functional Connectivity with Posterior Regions Tested in Experiment 2

Region of Interest	Contrast	MNI Coordinates		Surface area (mm^2)		Correlation with pVis			Correlation with pAud		
		Mean	SD	Mean	SD	r	t	p	r	t	p
Left Hemisphere											
sPCS	V > A	−33, −7, 47	6, 3, 5	296	133	0.47	4.54	0.01	−0.04	−0.29	1.00
iPCS	V > A	−44, 1, 33	6, 4, 6	206	118	0.50	5.77	0.003	−0.02	−0.19	1.00
tgPCS	A > V	−47, −5, 44	5, 5, 2	296	135	−0.04	−0.52	1.00	0.44	6.09	0.002
cIFS	A > V	−44, 12, 20	6, 4, 5	465	361	−0.04	−1.28	0.94	0.44	5.75	0.003
Right Hemisphere											
sPCS	V > A	34, −6, 48	6, 3, 6	475	177	0.58	6.06	0.002	0.02	0.22	1.00
iPCS	V > A	46, 3, 30	6, 3, 4	322	202	0.62	5.93	0.002	−0.11	−1.07	0.94
tgPCS	A > V	51, −4, 41	5, 1, 4	236	173	−0.11	−2.18	0.24	0.38	9.36	0.0001
cIFS	A > V	46, 20, 18	5, 7, 4	239	151	−0.04	−1.05	0.94	0.42	4.48	0.01

Regions are listed with their sensory bias: vision greater than audition (V > A) or audition greater than vision (A > V). Correlations between frontal and posterior regions are reported with the Pearson correlation (r), t-statistic (t), and p value (p) after Holm-Bonferroni correction for multiple comparisons.

the frontal ROIs defined by a visual-attention bias, sPCS and iPCS, showed a strong correlation with the posterior visual attention region, pVis, (white bars; all r > 0.4, p < 0.01, Holm-Bonferonni corrected), but no correlation with the posterior auditory attention region, pAud (black bars; all r < 0.05, p > 0.3, uncorrected). Conversely, in both hemispheres the frontal ROIs defined by an auditory-attention bias, tgPCS and cIFS, showed no positive correlation with the visual attention region, pVis, (all r < 0, p > 0.2, uncorrected, except right tgPCS: negative correlation p < 0.06, uncorrected) and a strong positive correlation with the auditory attention region, pAud (all r > 0.35, p < 0.05, Holm-Bonferonni corrected). Additionally, the correlations of each frontal ROI with pVis were significantly different from the correlations with pAud (all p < 0.02, Holm-Bonferonni corrected). The sensory-biased pattern in functional connectivity was observed across hemispheres and throughout the two networks (see Figure 3C). pVis and pAud were not correlated with each other in the left (r = 0.02) or right (r = −0.003) hemispheres. Shifting the statistical threshold used to define the frontal ROIs did not qualitatively change the correlations with posterior regions (see Figure S3A); neither did excluding Heschl's Gyrus when defining the pAud ROIs (see Figure S3B). Group average connectivity maps revealed a similar, but somewhat blurred, pattern of connectivity (see Figure S4).

Hierarchical clustering of the functional connectivity distance (1–r) between ROIs demonstrated a consistent pattern of two independent networks, which were organized by the same sensory bias detected in Experiment 1 (see Figure 3C). Bootstrap verification indicated that the 12 ROIs were organized into the same two networks in 98.1% of the 1000 bootstraps. A high cophenetic correlation (0.92) between the clustering matrix and the original distance matrix indicated that the cluster tree accurately represented the original correlation matrix. Using a thresholded correlation matrix to create network graphs revealed the same two networks shown in the hierarchical clustering (see Figure S2D). Combined with the task-based results of Experiment 1, these resting-state functional connectivity findings demonstrate that interdigitated nodes of auditory attention and visual attention networks exist bilaterally in lateral frontal cortex.

Experiments 3 and 4: Sensory Modality and Information Domain

The critical test of the domain recruitment hypothesis is to investigate whether these frontal attention networks are flexibly recruited based on the information domain (spatial or temporal) of the task even if sensory information is restricted to the non-preferred modality. Our domain recruitment hypothesis predicts that temporally demanding visual tasks will recruit the lateral frontal auditory-biased attention network, and that spatially demanding auditory tasks will recruit the lateral frontal visual-biased attention network. We tested this hypothesis by manipulating the spatial and temporal informational domain demands within visual (Experiment 3) and auditory (Experiment 4) sensory modalities using a change detection short-term memory paradigm. In these tasks, participants evaluated whether a target and a probe were the same (50% chance) or different (see Figures 4A and 5A, and Experimental Procedures).

In the visual tasks of Experiment 3 (Figure 4), participants either attempted to detect a change in orientation in one of the four simultaneously presented red bars (spatial task) or attempted to detect a change in the onset-timing pattern of the four sequentially presented red bars (temporal task). Subject performance was not significantly different between the two tasks (spatial, 81% ± 9%; temporal, 80% ± 5%; $t_8 = 0.13$, p = 0.90). The fMRI results demonstrate that the visual temporal task, but not the visual spatial task, recruited tgPCS and cIFS, the frontal regions of the auditory-biased attention network identified from Experiments 1 and 2. An ANOVA revealed an interaction between information domain and ROI within the visual modality ($F_{3,24} = 68.48$, p = 6.57e–12), but no main effect of hemisphere ($F_{1,8} = 0.001$, p = 0.98) or interactions between ROI and hemisphere ($F_{3,24} = 0.49$, p = 0.70), information domain and hemisphere ($F_{1,8} = 1.70$, p = 0.23), or ROI and information domain and hemisphere ($F_{3,24} = 1.22$, p = 0.33). We therefore combined ROIs from the two hemispheres. In the auditory-biased tgPCS and cIFS, the visual temporal task showed a stronger response than the spatial task, and only the visual temporal task, but not the visual spatial task, showed a significant BOLD response relative to the sensorimotor

A

Left Hemisphere ROIs

B

Right Hemisphere ROIs

C
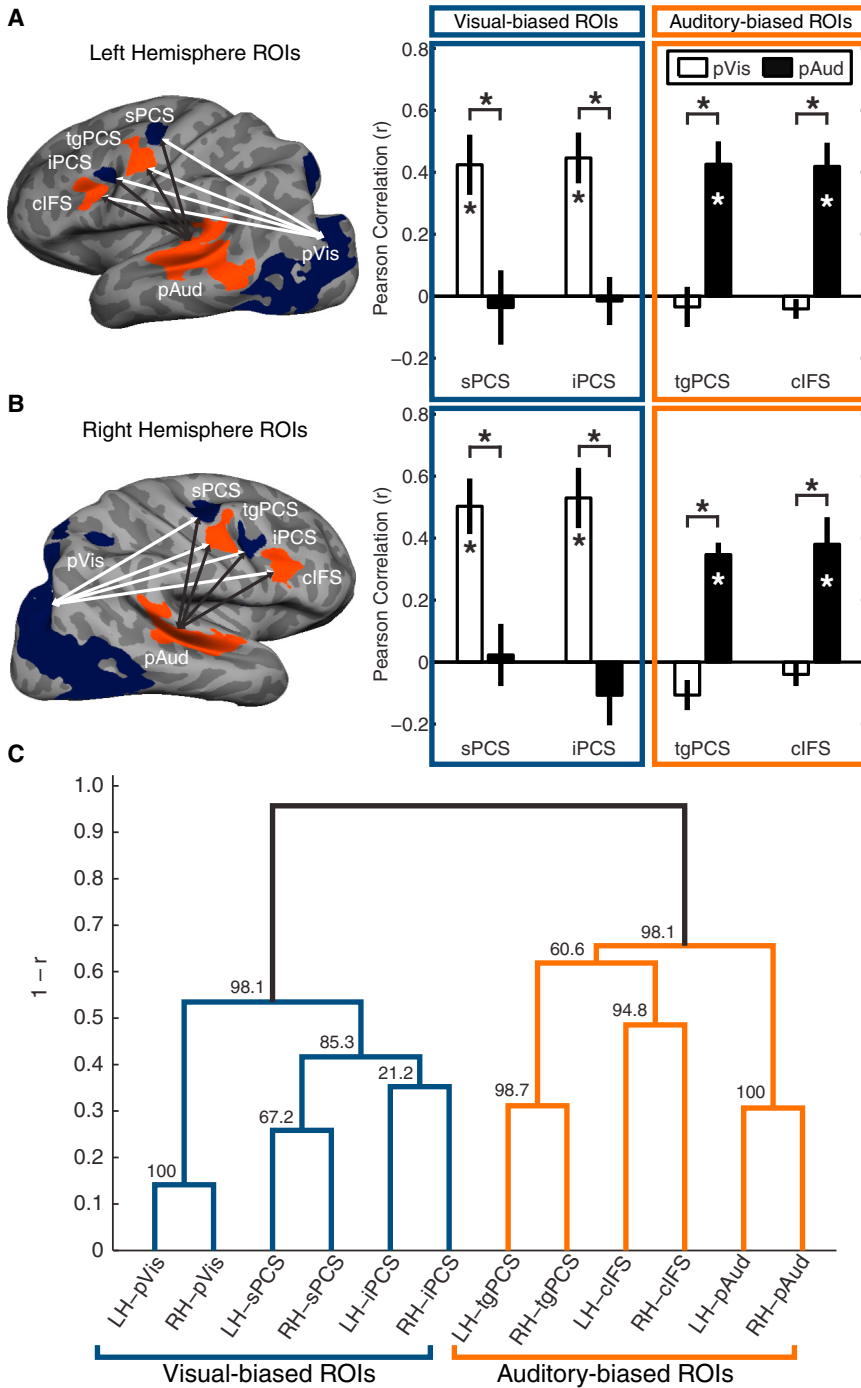
Figure 3. Group-Averaged Intrinsic Functional Connectivity in Experiment 2

(A and B) Within-hemisphere (left, right) average of Pearson correlation between resting-state time courses of each frontal ROI and the visual-biased pVis (white) and auditory-biased pAud (black) ROIs showing planned statistical comparisons. Brain images show the ROIs from one subject and illustrate which functional connectivity relationships were tested. Frontal ROIs are grouped (blue and orange boxes) by sensory-bias demonstrated in Experiment 1. The intrinsic functional connectivity pattern confirms the key finding of Experiment 1 that interleaved frontal areas participate in attentional networks strongly biased for sensory modality. Mean correlation from nine subjects; error bars reflect SEM.
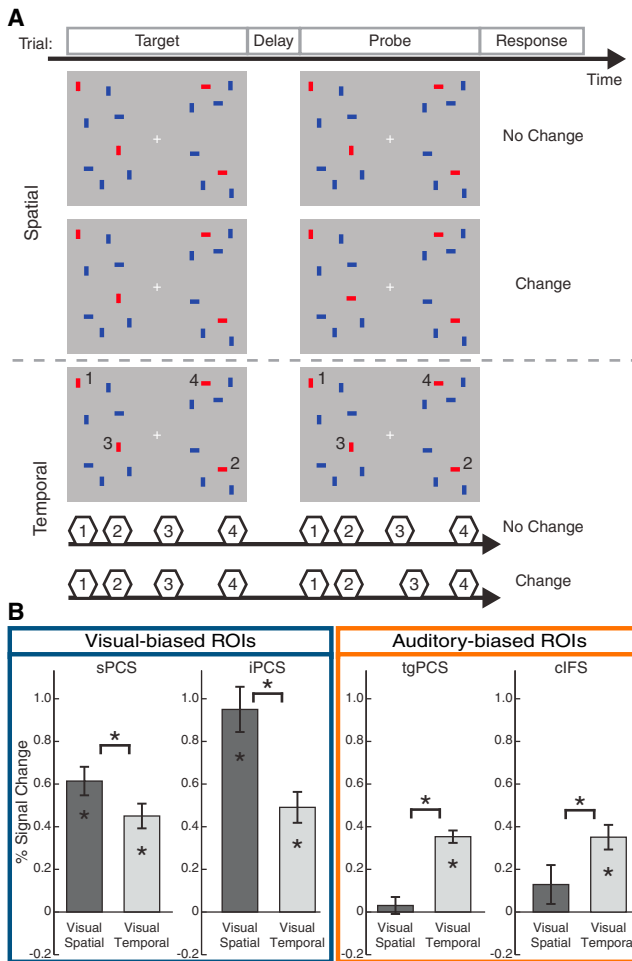
(C) Hierarchical cluster tree based on the (1−r) distance matrix between all 12 ROIs. Values in black indicate confidence based on percentage of 1,000 bootstrap average matrices matching each subtree.

tgPCS and cIFS, supporting the domain recruitment hypothesis.

In the auditory tasks of Experiment 4 (Figure 5), participants attempted to detect a change in the spatial location (spatial task) or onset-timing pattern (temporal task) of four sequentially presented complex tones. Although behavioral data was not significantly different between the two auditory tasks, there is a trend toward the temporal task being more difficult (spatial, 77% ± 12%; temporal, 67% ± 11%; $t_8 = 1.9$, $p = 0.09$); thus, recruitment of visual areas during the spatial task cannot be attributed to differences in task difficulty. In the fMRI results, we observed a complementary relationship to that seen in the visual tasks; high spatial demands in the auditory tasks flexibly recruited the visual-biased ROIs. An ANOVA revealed an interaction between information domain and ROI ($F_{3,8} = 12.78$, $p = 0.007$), but no main effect ($F_{1,8} = 0.64$, $p = 0.45$) or interactions with hemisphere (hemisphere*ROI: $F_{3,24} = 2.07$, $p = 0.13$; hemisphere*information domain: $F_{1,8} = 1.85$, $p = 0.21$; hemisphere*ROI*information domain: $F_{3,24} = 0.74$, $p = 0.42$); therefore, we again combined the two hemispheres for further analysis of the ROIs. Notably, the auditory spatial task showed stronger recruitment of visual-biased ROIs, sPCS and iPCS, compared to the temporal task (see Table 2). In the auditory-biased ROIs, tgPCS and cIFS, no differences in BOLD response were found between the auditory spatial and temporal tasks, with both tasks showing significant activation versus the sensorimotor control task. Both tasks also showed a significant

control (passive viewing + button press; see Experimental Procedures and Table 2 for details). Conversely, for the visual-biased ROIs, the visual spatial task showed greater BOLD response in sPCS and iPCS compared to the visual temporal task, and both tasks showed a significant response relative to sensorimotor control. Using a fixation baseline did not qualitatively change our results (see Figure S5). These results demonstrate that a purely visual task with high temporal demands can flexibly recruit the auditory-attention biased frontal regions,

A

Table 2. Statistical Results for Paired t Tests in Experiments 3 and 4

	Spatial		Temporal		Spatial versus Temporal	
	t	p	t	p	t	p
Experiment 3: Visual						
sPCS	**9.19**	**0.0001**	**7.77**	**0.0003**	**3.12**	**0.014**
iPCS	**8.95**	**0.0001**	**6.76**	**0.0006**	**8.06**	**0.0002**
tgPCS	0.77	0.46	**12.1**	**1.6e–5**	**-5.91**	**0.001**
cIFS	1.41	0.39	**6.06**	**0.0009**	**-4.21**	**0.006**
Experiment 4: Auditory						
sPCS	**9.70**	**7.4e-5**	**3.27**	**0.01**	**6.46**	**0.0008**
iPCS	**6.12**	**0.001**	**3.78**	**0.01**	**4.75**	**0.004**
tgPCS	**5.65**	**0.001**	**10.08**	**6.38e–5**	–0.65	1.0
cIFS	**6.05**	**0.001**	**6.31**	**0.001**	0.66	1.0

Values are t-statistic (t) and p value (p) after Holm-Bonferroni correction for multiple comparisons. Bold font indicates statistical significance (p < 0.05 corrected). Degrees of freedom = 8 for all tests.

B

Figure 4. Visual Spatial and Visual Temporal Change-Detection Tasks in Experiment 3

(A) Schematic of a single trial within a blocked design. Each trial began with a target stimulus (200 ms for spatial, 1,333 ms for temporal), followed by a 900 ms delay, and then a probe stimulus (2,300 ms for spatial, 1,333 ms for temporal). Subjects indicated "change" or "no change" along the attended dimension with a right hand button press. In the spatial task, subjects attempted to detect a change in the orientation of simultaneously presented red bars (see bottom left quadrant in "change" example). In the temporal task, stimuli were presented sequentially, indicated by digits (for illustration only), with variable times between the onsets of each bar (illustrated by spacing of hexagons).

(B) Average percent signal change (n = 9) relative to sensorimotor control within each frontal ROI (data combined across hemispheres) for the spatial (dark gray) and temporal (light gray) visual tasks. Error bars reflect SEM. Note the recruitment of the auditory-biased ROIs in the visual temporal task but not in the visual spatial task.

response versus sensorimotor control in sPCS and iPCS. Using a fixation baseline did not qualitatively change our results (see Figure S6). Although sPCS and iPCS can be driven by eye movements (e.g., Paus 1996; Corbetta et al., 1998), the observed functional differences cannot be attributed to eye movements or motor responses: eye-tracking during the auditory task revealed no difference in the number of eye movements between

the spatial and temporal task ($t_6 = 0.35$, p = 0.74, see Figure S7 and Supplemental Experimental Methods) and motor responses were also equivalent across tasks. As a final analysis we combined the results from Experiments 3 and 4 into a single three-way ANOVA and observed a highly significant 3-way interaction between ROI, sensory modality, and information domain ($F_{3,24} = 60.02$, p = 2.64e−11). Taken together, the increased response for the visual temporal compared to the visual spatial task in auditory-biased frontal ROIs and the increased response for the auditory spatial compared to auditory temporal task in visual-biased frontal ROIs strongly support the domain recruitment hypothesis.

DISCUSSION

Experiments 1 and 2 demonstrate that sensory modality is a key factor in the functional organization of four regions in human lateral frontal cortex, while Experiments 3 and 4 provide critical tests supporting the domain recruitment hypothesis. Four functionally distinct, anatomically interdigitated regions run from the intersection of the precentral sulcus and the superior frontal sulcus down the precentral sulcus and into the caudal inferior frontal sulcus. The two visual-biased attention network areas that we identify are located in the sPCS and iPCS, while the two auditory-biased network areas lie in adjacent cortex, where the transverse gyrus intersects with the precentral sulcus (tgPCS) and just anterior to iPCS in the caudal IFS (cIFS). The sensory biases of these frontal cortical regions are demonstrated by (1) a direct contrast of activation during auditory attention versus visual attention in a task with matched spatial and temporal demands and (2) highly selective intrinsic functional connectivity (resting-state) with posterior cortical areas with known sensory biases. Consistent with the domain recruitment hypothesis, Experiments 3 and 4 demonstrate that both areas of each network can be flexibly recruited by the nonpreferred sensory modality if the information demands of the task play to the strength (i.e., spatial or temporal information) of the sensory modality associated with a particular

A

B
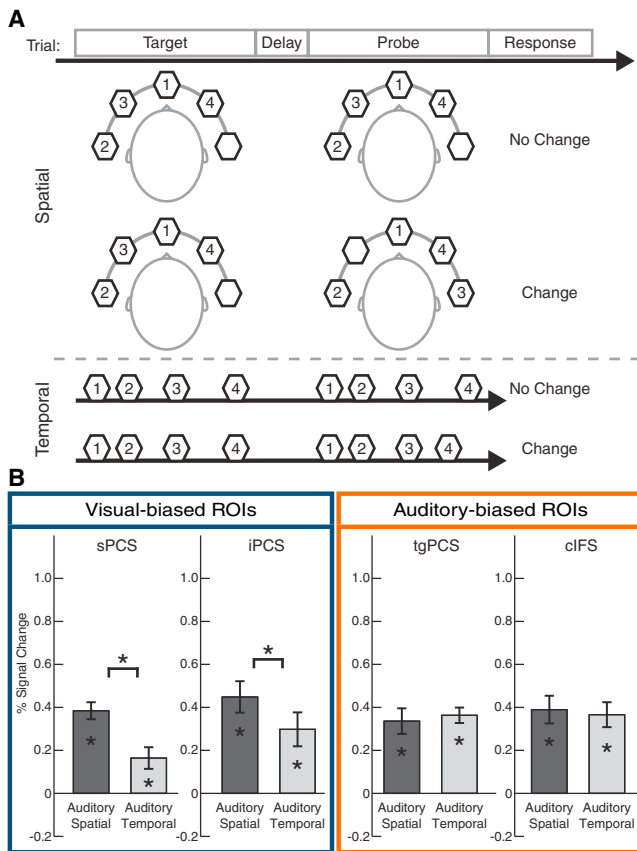
Figure 5. Auditory Spatial and Auditory Temporal Change-Detection Tasks in Experiment 4

(A) Schematic of a single trial within a blocked design. Each trial comprised a target stimulus (2,350 ms), followed by a 900 ms delay, and then a probe stimulus (2,350 ms). Subjects indicated "change" or "no change" along the attended dimension with a right hand button press. Within each stimulus presentation, complex spatialized tones were presented sequentially. In the spatial task, subjects attempted to detect a change in the location of any of the tones. In the temporal task, subjects tried to detect changes in the onset-timing pattern of the tones.

(B) Average percent signal change (n = 9) relative to sensorimotor control within each frontal ROI (data combined across hemispheres) for the spatial (dark gray) and temporal (light gray) auditory tasks. Error bars reflect SEM. Note the stronger recruitment of visual-biased ROIs for the auditory spatial task compared to the auditory temporal task.

region. A purely visual task with high temporal demands recruited the auditory-biased regions, tgPCS and cIFS, while a purely auditory task with high spatial demands recruited the visual-biased regions, sPCS and iPCS. Our findings reveal two distinct attentional networks that are strongly linked with different sensory modalities (vision or audition) and are also strongly linked with different information domain representations (space or time, respectively).

Our findings in the visual-biased regions are consistent with prior studies showing strong recruitment in these areas during visual attention and short-term memory tasks (Paus, 1996; Courtney et al., 1998; Hagler and Sereno, 2006; Kastner et al., 2007; Jerde et al., 2012). Consistent with a prior fMRI study (Tark

and Curtis, 2009), we found that auditory spatial attention recruits the sPCS. In addition, we observed recruitment of iPCS. The flexible recruitment of iPCS and sPCS—the putative human homolog of nonhuman primate frontal eye field (FEF)—cannot be attributed to eye motor response, as we found no differences in eye movements between the spatial and temporal auditory tasks. Prior visual fMRI studies using spatial attention, spatial working memory, and/or spatial motor intention (e.g., saccade mapping) have identified visual topographic maps in the vicinity of sPCS and iPCS with a gap between the two regions (Hagler and Sereno, 2006; Kastner et al., 2007; Jerde et al., 2012). This link between frontal visual areas, sPCS and iPCS, and spatial processing is central to one key aspect of the domain recruitment hypothesis. To account for the complementary aspect of the hypothesis, we conjecture that specialized representations for timing and rhythm exist in the frontal auditory-biased regions, tgPCS and cIFS. This conjecture is supported by neuroimaging work indicating that perception and rehearsal of rhythms in the absence of overt movements drives activation within lateral frontal cortex (Karabanov et al., 2009; Chapin et al., 2010).

In this study, tgPCS and cIFS demonstrate a clear bias for auditory attention but can be flexibly recruited under high temporal demands. Here, we introduced nomenclature for tgPCS; however, prior studies have reported auditory task activation in the broad vicinity of tgPCS, bilaterally, for pitch or tonal memory (Gaab et al., 2003; Koelsch et al., 2009) and verbal memory (Koelsch et al., 2009). Auditory task activity in the vicinity of cIFS, though typically reported on the inferior frontal gyrus, has previously been identified in the right hemisphere for working memory of voices (Rämä and Courtney, 2005) and attention to tones (Braga et al., 2013) and in the left hemisphere for verbal working memory (Awh et al., 1996; Crottaz-Herbette et al., 2004) and attention to pitch (Hill and Miller, 2010). cIFS is distinct from Broca's area, as Broca's area lies ventral to the IFS. Postmortem receptor mapping has revealed fine-scale anatomical subdivisions in this vicinity (Amunts et al., 2010), but further study investigating how the presently defined functional areas relate to those anatomical definitions is needed.

We observed that sensory modality is a key factor in the functional organization of caudal lateral frontal cortex; in contrast, there is an extensive literature positing that a domain-general, task-positive, or multiple demand network exists in lateral frontal cortex (Duncan and Owen, 2000; Fox et al., 2005; Duncan, 2010). Most prior human neuroimaging work has reported that auditory and visual responses merge in frontal cortex (Lewis et al., 2000; Johnson and Zatorre, 2006; Ivanoff et al., 2009; Karabanov et al., 2009; Tark and Curtis, 2009; Tombu et al., 2011; Braga et al., 2013). Although some studies have reported a bias for one modality, these reports generally cannot exclude task or task difficulty biases as the source of the sensory bias (Crottaz-Herbette et al., 2004; Jantzen et al., 2005; Rämä and Courtney, 2005; Sallet et al., 2013). Here, we clearly demonstrate distinct regions of lateral frontal cortex that are biased for attention to sensory modality. However, our findings do not rule out the existence of some domain-general processing elements within lateral frontal cortex; conceivably, several functional organizations are multiplexed within this region of cortex. Nevertheless, our results demonstrate that the multiple-demand view is an incomplete

description that overlooks the important role of sensory modality in the functional organization of lateral frontal cortex. By analyzing data from individual subjects on their cortical surfaces, we were able to obtain a higher effective spatial resolution than is typically obtained with group-averaging methods. These methods may have been critical to resolving multiple distinct visual-biased and auditory-biased attention regions where prior studies found responses independent of sensory modality. Consistent with our findings, a recent multivariate analysis study indicated that posterior lateral frontal cortex contains information about sensory modality, but this study did not identify specific visual-biased and auditory-biased frontal cortical areas (Tamber-Rosenau et al., 2013).

Our findings are largely orthogonal to reports of hierarchical organization in the LFC (e.g., Koechlin et al., 2003; Badre et al., 2010); however, we note that the two most caudal regions in these studies (i.e., PMD and pre-PMD) may align with sPCS and iPCS. Similar coordinates have been reported in studies of cognitive control (Brass et al., 2005) and salience detection (Corbetta and Shulman, 2002). Future studies will be needed to investigate their colocalization as well as the role of sensory modality in relation to the proposed hierarchical organization of frontal cortex.

The domain recruitment hypothesis is a neural hypothesis related to the modality appropriateness hypothesis, a perceptual hypothesis that describes the biased relationships among vision and audition and space and time when conflicting sensory information arises (cf. Alais and Burr, 2004, for important exceptions). The domain recruitment hypothesis extends this concept to neural responses under higher cognitive demands. Several prior behavioral studies investigating short-term memory for spatial and/or temporal information presented in visual and/or auditory modalities have reported that the visual modality is superior for spatial STM and that the auditory modality is superior for temporal STM (e.g., Balch and Muscatelli, 1986; Glenberg et al., 1989; Collier and Logan, 2000; Guttman et al., 2005; McAuley and Henry, 2010). Cross-modal recoding (e.g., hearing visual rhythms) may occur when the information domain of the task is not "appropriate" to the native stimulus modality; however, debate remains as to whether such recoding is automatic and obligatory or controlled and strategic (Guttman et al., 2005; McAuley and Henry, 2010). Several subjects in the present study reported that they could "visualize" the auditory spatial locations and/or "hear" the visual rhythms, thus the neural domain recruitment observed here likely reflects a form of cross-sensory STM recoding.

It is instructive to contrast our domain recruitment hypothesis with the well-known domain specificity hypothesis, which argues that working memory processes that are specific to an information domain—object identity or spatial location ("what versus where")—may be anatomically localized in PFC (Goldman-Rakic, 1996). Although the validity of this hypothesis has been debated (Romanski, 2007; Rao et al., 1997; Postle et al., 2000), it should be noted that the domain recruitment hypothesis differs in three primary ways: (1) it addresses temporal versus spatial processing ("when versus where"), (2) it suggests that information domains are biased toward sensory modalities, and (3) it proposes that cortical regions can be flexibly recruited. The

domain recruitment hypothesis predicts biases for both information domain and sensory modality in cortical regions.

We observed asymmetries between visual and auditory processing within these frontal regions. In Experiment 1, the auditory-biased network regions were not driven by visual spatial attention (versus sensorimotor baseline), but the visual-biased network regions were driven by auditory spatial attention. Given the spatial nature of the task, this result is predicted by the domain recruitment hypothesis. In both Experiment 3 (visual stimuli) and Experiment 4 (auditory stimuli), the visual-biased network regions were more strongly activated in the spatial than the temporal tasks. In contrast, the auditory-biased network regions were more strongly activated in the temporal visual than spatial visual task, but were strongly activated in both auditory tasks of Experiment 4. This asymmetry (i.e., visual-biased regions showed a difference between visual spatial and visual temporal STM, while auditory-biased regions showed no difference between auditory spatial and auditory temporal STM) is not central to the domain recruitment hypothesis, but it is also not predicted by the hypothesis. One possible explanation is that auditory rhythms were encoded in both the spatial and temporal forms of the task.

Our findings reconcile an apparent discrepancy between the human and nonhuman primate literature regarding the functional organization of caudal lateral frontal cortex. Studies in nonhuman primates have indicated sensory-biased regions of caudal prefrontal cortex (Petrides and Pandya, 1999; Romanski and Goldman-Rakic, 2002; Romanski, 2007). Experiments 1 and 2 clearly demonstrate that human lateral frontal cortex also exhibits functional divisions organized around sensory modality regions. In nonhuman primates, the organization from dorsal to ventral appears to run auditory (BA8b), visual (BA8a), visual (BA45), auditory (BA12/47). We observed a different dorsal to ventral organization: visual (sPCS), auditory (tgPCS), visual (iPCS), and auditory (cIFS). In humans, the precentral sulcus is interrupted by a transverse gyrus (tgPCS) (Ono et al., 1990). In nonhuman primates, the arcuate sulcus, which serves as the caudal border of the sensory areas, is unbroken. This gross anatomical difference corresponds to the location of a key difference in functional organization: an auditory region between two visual regions. Future research in humans and nonhuman primates will help to elucidate the organization of the frontal lobe, and the findings presented here represent a significant step in defining and understanding the functional roles of distinct networks in human lateral frontal cortex.

EXPERIMENTAL PROCEDURES

Eleven healthy individuals (mean age, 27.1 years; range, 22–31 years, 5 females) participated in the experiments. All participants were right-handed, native English speakers with normal or corrected-to-normal vision, received monetary compensation, and gave informed consent to engage in the study according to the procedures approved by the Institutional Review Board at Boston University and/or Partners Healthcare. Participants were required to hold gaze at a central fixation point in all experiments and were trained to hold fixation. One participant was excluded from all analysis because of head movements. A second participant was excluded from Experiments 2–4 due to difficulties in defining ROIs on the basis of the results of Experiment 1. Two authors (S.W.M. and M.L.R.) participated as subjects.

Data Collection

Each subject participated in a minimum of five sets of scans across multiple sessions and separate behavioral training sessions. In addition to the four fMRI experiments, high-resolution structural scans were collected to support anatomical reconstruction of the cortical hemispheric surfaces. Imaging was performed at the Center for Brain Science Neuroimaging Facility at Harvard University on a 3-T Siemens Tim Trio scanner with a 32-channel matrix coil. A high-resolution (1.0 × 1.0 × 1.3 mm) magnetization-prepared rapid gradient-echo sampling structural scan was acquired for each subject. The cortical surface of each hemisphere was computationally reconstructed from this anatomical volume using FreeSurfer software (http://surfer.nmr.mgh.harvard.edu/). For functional studies, T2*-weighted gradient echo, echo-planar images were collected using 42 3 mm slices (0% skip), oriented axially (time echo 30 ms, time repetition [TR] 2,600 ms, in-plane resolution 3.125 × 3.125 mm). In the visual spatial task, 7 of 11 subjects were scanned on an identically equipped Siemens Tim Trio scanner at the Martinos Center for Biomedical Imaging at Massachusetts General Hospital.

Stimulus Presentation

Visual stimuli were presented using a liquid crystal display projector illuminating a screen within the scanner bore. The display extended across a visual angle of ∼14° radius horizontally and ∼11° radius vertically. The audio system (Sensimetrics, http://www.sens.com) included an audio amplifier, S14 transformer, and MR-compatible earphones. Inside the MR scanner, subject responses were collected using an MR-compatible button box.

Experiment 1: Sustained Visual and Auditory Attention

Participants monitored one of four (two auditory, two visual) rapid serial streams of distractor letters ("A," "F," "G," "H," "J," "K," "L," "M," "N," "P," "R," "X," and "Y") for the presentation of any digit (1–4), while ignoring the other streams containing only digits (1–9, excluding the two-syllable digit 7; see Figure 1 for example). At the beginning of each block, participants were directed by an audiovisual cue to attend to one of the four streams ("watch left," "watch right," "listen left," "listen right"), perform a sensorimotor control ("passive"), or simply hold fixation with only a central cross presented ("fixation"). Participants were instructed to press the corresponding button (1–4) whenever a digit was presented in the attended stream (3 times per 26 s block). Ten stimuli (2 auditory, 2 visual, and 6 visual flankers) were simultaneously presented for 300 ms followed by a 350 ms interstimulus interval (ISI). Each participant completed 3–6 fMRI runs, with each run containing 12 blocks evenly divided into 6 conditions: attend to left auditory, attend to right auditory, attend to left visual, attend to right visual, sensorimotor control, and fixation. Each block lasted 26 s, included 40 serial stimulus presentations, and was preceded by a 2.6 s cue period (voice and text indicating the next block). In the sensorimotor control condition, all streams contained only digits and participants were instructed to press each of the 4 available buttons 1 time at a relaxed pace at any point during the block.

The visual stimuli (white, 1.5° × 1.5°, presented on a dark gray background) were located 4.5° to the left and right of a central fixation cross (1.5° × 1.5°) and were flanked by three additional distractor streams on each side that always contained distractor digits. Auditory streams were generated from monaural recordings of 8 digits and 13 letters spoken by a single male talker. Each digit/letter was sampled at 44.1 kHz with a duration of 300 ms and windowed with cosine-squared onset and offset ramps (30 ms ramp time). Each monaural digit recording was used to generate a binaural, lateralized signal in which the signal at the 2 ears was identical, except for an interaural time delay (ITD) of 800 μs leading either right or left (with no interaural level difference). This manipulation resulted in lateralized percepts, with the digits perceived as coming from either right or left of the median plane, depending on the sign of the ITD.

Three subjects participated in an additional event-related sustained attention task using a similar stimulus set. These data were not included in the data for Experiment 1, but rather served as a back-up method for defining frontal ROIs in three individual subjects for use in Experiments 2–4. In the event-related task, participants attended to one of four streams of letters (two auditory, two visual), while ignoring all other streams. Each of the stimulus streams was assigned a digit 1–4. A digit presented in the attended stream indicated that the participant should either shift their attention to a new stream

or continue to maintain attention to the current stream (if the presented digit matched the currently attended stream). In the event-related task, the stimuli in the two visual streams (no flankers) were presented centrally.

Experiment 2: Intrinsic Functional Connectivity

Subjects also participated in resting-state scans, in which participants were instructed to keep their eyes open, maintain fixation on a centrally presented cross, allow their minds to wander, and avoid repetitive activities such as counting. Each run was either 139 or 256 time points, and subjects participated in one to two runs. Imaging parameters were the same as in Experiment 1.

Experiment 3: Spatial and Temporal Visual Tasks

Both Experiment 3 (visual) and Experiment 4 (auditory) manipulated the information domain (spatial or temporal) demands of the task within a sensory modality. All four tasks used a change detection paradigm where each trial comprised a target stimulus, followed by a 900 ms delay with only a fixation cross, and then a probe (50% chance of change from target in the attended feature) and response period. Subjects were instructed to respond using a right hand button press to denote whether the attended feature changed between the target and probe stimulus presentation. Each task was compared to a sensorimotor control condition, where the stimuli matched the active task condition and subjects were instructed to refrain from doing the task but to respond with a random button press at the end of each trial. Each run was divided into blocks of task and sensorimotor control conditions. The two visual tasks occurred in different runs and in different sessions for 8 of 11 of subjects. Imaging parameters were the same as in the prior experiments.

In the visual spatial task (see Figure 4A), participants were instructed to covertly attend to the orientation of four red colored bars, oriented vertically or horizontally and presented among 12 blue distractor bars. Bars were evenly distributed across hemifields with two red bars in each hemifield. Each bar subtended 0.3° × 0.9° of visual angle. The target stimulus was presented for 200 ms, and the probe and response period was 2,300 ms (probe stimulus on for 1,900 ms). In the "change" trials, the orientation of one of the four red target bars would change by 90° between the target and probe stimulus periods. In the sensorimotor control condition, all bars were blue and no change occurred between the target and probe stimuli.

In the visual temporal task (see Figure 4A), participants attended to the onset-timing pattern of the red bars. Both the target stimulus and the probe stimulus were presented in 1.33 s periods, beginning with the onset of 12 blue bars and followed by the sequential onset of four red bars. The stimulus onset asynchrony (SOA) of the bars ranged between 133 and 400 ms. Within a trial, the bars always appeared in the same orientation, location, and order. In "change" trials, the timing pattern of the probe stimuli differed from that of the target stimuli. In the sensorimotor control condition, all bars were blue and no change occurred between the target and probe stimuli.

Experiment 4: Spatial and Temporal Auditory Tasks

Experiment 4 (see Figure 5A) used a change detection paradigm, mirroring Experiment 3. The same auditory stimuli were used in the spatial and temporal tasks (only the attended feature changed). Each stimulus comprised a sequence of four complex tones presented over a 2,350 ms period, with each tone including frequencies in the first three harmonics of three fundamental frequencies (130.81 Hz, 174.61 Hz, and 207.65 Hz) at equal intensity, ramped on and off with a 16 ms-long cosine squared window. Each tone lasted 370 ms with between tone intervals (BTIs) ranging from 120 to 420 ms. All tones were the same combination of sine waves and were separated by irregular BTIs. ITDs of –1,000 μs, –350 μs, 0 μs, 350 μs, and 1,000 μs were used to spatially localize the tones along the azimuth. The first tone in the sequence was always located centrally (0 μs ITD). In the auditory spatial task, subjects attended to the locations of the tones. In "change" trials, one of the three tones following the initial centered tone was relocated to the unused spatial location. In the temporal task, subjects attended to the timing pattern of the sequence of tones. In "change" trials, one of the BTIs changed by at least 50 ms between the target and probe stimulus. In both of the auditory tasks, the other dimension (location or timing) was the same for the target and probe stimuli. In the sensorimotor control condition, no change occurred between the target and probe stimuli along the dimension of either timing or location.

Eye Tracking

Participants were eye-tracked in the scanner in Experiments 1 and 4; see Supplemental Information for details.

fMRI Analysis

Functional data were analyzed using Freesurfer/FS-FAST (CorTech, Inc.) with an emphasis on localizing distinct cortical areas on individual subject's cortical surfaces. All analysis was performed on subject-specific anatomy. All subject data were registered to the individual's anatomical data using the mean of the functional data, motion corrected by run, slice-time corrected, intensity normalized, resampled onto the individual's cortical surface (voxels to vertices), and spatially smoothed on the surface with a 3 mm full-width half-maximum Gaussian kernel.

Analysis of the Experiment 1, 3, and 4 scans used standard procedures and Freesurfer FS-FAST software (Version 5.1.0). Scan time series were analyzed vertex-by-vertex on the surface using a general linear model (GLM) whose regressors matched the time course of the experimental conditions. The time points of the cue period were excluded by assigning them to a regressor of no interest. In addition, singular value decomposition reduced the six vectors from motion correction (degrees of freedom) to 3 eigenvectors, which were included as nuisance regressors. The canonical hemodynamic response function was convolved with the regressors before fitting; this canonical response was modeled by a γ function with a delay of $\delta = 2.25$ s and decay time constant of $\tau = 1.25$. A contrast between different conditions produced t-statistics for each vertex for each subject.

In Experiment 1, ROIs were defined on each individual subject based on a direct contrast of blocks in which the subject attended to one of the auditory streams and blocks in which the subject attended to one of the visual streams. This direct contrast was liberally thresholded at $p < 0.05$ uncorrected to maximally capture vertices showing a bias for attention to either the auditory or the visual stimuli (this resulted in all ROIs being larger than 48 mm^2). All behavioral data were compared using two-tailed paired t tests across conditions.

For ROI analysis in Experiments 1, 3 and 4, the percentage signal change data were extracted for all voxels in the ROI and averaged across all blocks for all runs for each condition. The percent signal change measure was defined relative to the average activation level during the sensorimotor control condition. Separately for Experiments 3 and 4, we evaluated the ROI data extracted for each subject to test the relationship between the factors of ROI (sPCS, tgPCS, iPCS, cIFS), hemisphere (left, right), and information domain (spatial, temporal) using repeated-measures analysis of variance (ANOVA) in SPSS (http://www.ibm.com/software/analytics/spss/). If Mauchly's test indicated a violation of sphericity (e.g., Experiment 4), lower bound corrections were applied to the degrees of freedom of the F-test to reduce the likelihood of false positives in the ANOVA. When no interaction involving hemisphere was found, we combined ROI data across hemispheres. Based on our hypotheses, we were primarily interested in interactions between ROI and task. When this interaction was significant in the ANOVA, we conducted a two-tailed paired t test for each ROI to test the effect of information domain (four comparisons). Within each experiment, the p values from these t tests were corrected for multiple comparisons using the Holm-Bonferroni method. Additional paired t tests were performed (and similarly corrected) to test if each task was significantly activated in each frontal ROI relative to its sensorimotor control.

In Experiment 2, the resting-state data underwent additional processing using Matlab to reduce artifacts that could lead to spurious functional connectivity. Following the preprocessing described above, the data underwent multiple regression with nuisance regressors including the average white matter signal, average signal from the ventricular regions of interest, whole brain signal averaged across the whole brain, and 12 motion regressors (6 motion parameters from Freesurfer motion correction and their 6 temporal derivatives). We removed motion time points and applied a band-pass filter with $0.01 < f < 0.08$ Hz. We then calculated the average time course within each of the 12 ROIs defined in Experiment 1 for each subject. The Pearson's correlation coefficients were calculated for each posterior ROI (pVis and pAud) with each frontal ROI (sPCS, tgPCS, iPCS, and cIFS) within each hemisphere for hypothesis-driven tests. Group-level significance of correlations was tested using t tests on the z values, but graphs show mean Pearson correlations. All t tests were then corrected for multiple comparisons using the Holm-Bonferroni

method. Hierarchical clustering was conducted using a distance measure of $(1-r)$ and a common average linkage method (UPGMA). Cluster tree branch points were validated using 1,000 bootstraps to calculate the percentage of bootstrap trees containing a subtree that matched a subtree in the original cluster tree. See Supplemental Information for Experiment 2 analysis details.

REFERENCES

Adams, J. (2006). Neuroanatomical considerations of speech processing. In Listening to Speech: An Auditory Perspective, S. Greenberg and W.A. Ainsworth, eds. (Mahwah, NJ: Lawrence Erlbaum Associates), pp. 79–90.

Agmon-Snir, H., Carr, C., and Rinzel, J. (1998). The role of dendrites in auditory coincidence detection. Nature 393, 268–272.

Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. Curr. Biol. 14, 257–262.

Amunts, K., Lenzen, M., Friederici, A., Schleicher, A., Morosan, P., Palomero-Gallagher, N., and Zilles, K. (2010). Broca's region: novel organizational principles and multiple receptor mapping. PLoS Biol. 8, http://dx.doi.org/10.1371/journal.pbio.1000489.

Awh, E., Jonides, J., Smith, E., Schumacher, E., Koeppe, R., and Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: evidence from positron emission tomography. Psychol. Sci. 7, 25–31.

Badre, D., Kayser, A., and D'Esposito, M. (2010). Frontal cortex and the discovery of abstract action rules. Neuron 66, 315–326.

Balch, W.R., and Muscatelli, D.L. (1986). The interaction of modality condition and presentation rate in short-term contour recognition. Percept. Psychophys. 40, 351–358.

Barbas, H., and Mesulam, M. (1981). Organization of afferent input to subdivisions of area 8 in the rhesus monkey. J. Comp. Neurol. 200, 407–431.

Braga, R., Wilson, L., Sharp, D., Wise, R., and Leech, R. (2013). Separable networks for top-down attention to auditory non-spatial and visuospatial modalities. Neuroimage 74, 77–86.

Brass, M., Derrfuss, J., Forstmann, B., and von Cramon, D.Y. (2005). The role of the inferior frontal junction area in cognitive control. Trends Cogn. Sci. 9, 314–316.

Chapin, H., Zanto, T., Jantzen, K., Kelso, S., Steinberg, F., and Large, E. (2010). Neural responses to complex auditory rhythms: the role of attending. Front. Psychol. 1, 224.

Collier, G.L., and Logan, G. (2000). Modality differences in short-term memory for rhythms. Mem. Cognit. 28, 529–538.

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

Corbetta, M., Akbudak, E., Conturo, T.E., Snyder, A.Z., Ollinger, J.M., Drury, H.A., Linenweber, M.R., Petersen, S.E., Raichle, M.E., Van Essen, D.C., and Shulman, G.L. (1998). A common network of functional areas for attention and eye movements. Neuron *21*, 761–773.

Courtney, S., Petit, L., Maisog, J., Ungerleider, L., and Haxby, J. (1998). An area specialized for spatial working memory in human frontal cortex. Science *279*, 1347–1351.

Crottaz-Herbette, S., Anagnoson, R., and Menon, V. (2004). Modality effects in verbal working memory: differential prefrontal and parietal responses to auditory and visual stimuli. Neuroimage *21*, 340–351.

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends Cogn. Sci. *14*, 172–179.

Duncan, J., and Owen, A. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends Neurosci. *23*, 475–483.

Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., and Raichle, M.E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc. Natl. Acad. Sci. USA *102*, 9673–9678.

Gaab, N., Gaser, C., Zaehle, T., Jancke, L., and Schlaug, G. (2003). Functional anatomy of pitch memory—an fMRI study with sparse temporal sampling. Neuroimage *19*, 1417–1426.

Gawne, T. (2000). The simultaneous coding of orientation and contrast in the responses of V1 complex cells. Exp. Brain Res. *133*, 293–302.

Glenberg, A.M., Mann, S., Altman, L., Forman, T., and Procise, S. (1989). Modality effects in the coding reproduction of rhythms. Mem. Cognit. *17*, 373–383.

Goldman-Rakic, P. (1996). Regional and cellular fractionation of working memory. Proc. Natl. Acad. Sci. USA *93*, 13473–13480.

Guttman, S., Gilroy, L., and Blake, R. (2005). Hearing what the eyes see: auditory encoding of visual temporal sequences. Psychol. Sci. *16*, 228–235.

Hagler, D., and Sereno, M. (2006). Spatial maps in frontal and prefrontal cortex. Neuroimage *29*, 567–577.

Hill, K., and Miller, L. (2010). Auditory attentional control and selection during cocktail party listening. Cereb. Cortex *20*, 583–590.

Ivanoff, J., Branning, P., and Marois, R. (2009). Mapping the pathways of information processing from sensation to action in four distinct sensorimotor tasks. Hum. Brain Mapp. *30*, 4167–4186.

Jantzen, K., Steinberg, F., and Kelso, J. (2005). Functional MRI reveals the existence of modality and coordination-dependent timing networks. Neuroimage *25*, 1031–1042.

Jerde, T.A., Merriam, E.P., Riggall, A.C., Hedges, J.H., and Curtis, C.E. (2012). Prioritized maps of space in human frontoparietal cortex. J. Neurosci. *32*, 17382–17390.

Johnson, J., and Zatorre, R. (2006). Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. Neuroimage *31*, 1673–1681.

Joris, P., Carney, L., Smith, P., and Yin, T. (1994). Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. J. Neurophysiol. *71*, 1022–1036.

Karabanov, A., Blom, O., Forsman, L., and Ullén, F. (2009). The dorsal auditory pathway is involved in performance of both visual and auditory rhythms. Neuroimage *44*, 480–488.

Kastner, S., DeSimone, K., Konen, C., Szczepanski, S., Weiner, K., and Schneider, K. (2007). Topographic maps in human frontal cortex revealed in memory-guided saccade and spatial working-memory tasks. J. Neurophysiol. *97*, 3494–3507.

Koechlin, E., Ody, C., and Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. Science *302*, 1181–1185.

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., and Gruber, O. (2009). Functional architecture of verbal and tonal working memory: an FMRI study. Hum. Brain Mapp. *30*, 859–873.

Lewis, J., Beauchamp, M., and DeYoe, E. (2000). A comparison of visual and auditory motion processing in human cerebral cortex. Cereb. Cortex *10*, 873–888.

McAuley, J.D., and Henry, M.J. (2010). Modality effects in rhythm processing: Auditory encoding of visual rhythms is neither obligatory nor automatic. Atten. Percept. Psychophys. *72*, 1377–1389.

O'Connor, N., and Hermelin, B. (1972). Seeing and hearing and space and space and time. Percept. Psychophys. *11*, 46–48.

Ono, M., Kubick, S., and Albernathey, C. (1990). Atlas of the Cerebral Sulci (New York: Thieme Medical Publishers).

Paus, T. (1996). Location and function of the human frontal eye-field: a selective review. Neuropsychologia *34*, 475–483.

Petrides, M., and Pandya, D. (1999). Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. Eur. J. Neurosci. *11*, 1011–1036.

Pick, H., Warren, D., and Hay, J. (1969). Sensory conflict in judgments of spatial direction. Percept. Psychophys. *6*, 203–205.

Postle, B., Stern, C., Rosen, B., and Corkin, S. (2000). An fMRI investigation of cortical contributions to spatial and nonspatial visual working memory. Neuroimage *11*, 409–423.

Rämä, P., and Courtney, S. (2005). Functional topography of working memory for face or voice identity. Neuroimage *24*, 224–234.

Rao, S., Rainer, G., and Miller, E. (1997). Integration of what and where in the primate prefrontal cortex. Science *276*, 821–824.

Recanzone, G. (2003). Auditory influences on visual temporal rate perception. J. Neurophysiol. *89*, 1078–1093.

Romanski, L. (2007). Representation and integration of auditory and visual stimuli in the primate ventral lateral prefrontal cortex. Cereb. Cortex *17*, i61–i69.

Romanski, L., and Goldman-Rakic, P. (2002). An auditory domain in primate prefrontal cortex. Nat. Neurosci. *5*, 15–16.

Sallet, J., Mars, R., Noonan, M., Neubert, F.-X., Jbabdi, S., O'Reilly, J., Filippini, N., Thomas, A.G., and Rushworth, M.F. (2013). The organization of dorsal frontal cortex in humans and macaques. J. Neurosci. *33*, 12255–12274.

Salmi, J., Rinne, T., Degerman, A., Salonen, O., and Alho, K. (2007). Orienting and maintenance of spatial attention in audition and vision: multimodal and modality-specific brain activations. Brain Struct. Funct. *212*, 181–194.

Shams, L., Kamitani, Y., and Shimojo, S. (2000). Illusions. What you see is what you hear. Nature *408*, 788.

Silver, M., and Kastner, S. (2009). Topographic maps in human frontal and parietal cortex. Trends Cogn. Sci. *13*, 488–495.

Swisher, J., Halko, M., Merabet, L., McMains, S., and Somers, D. (2007). Visual topography of human intraparietal sulcus. J. Neurosci. *27*, 5326–5337.

Tamber-Rosenau, B.J., Dux, P.E., Tombu, M.N., Asplund, C.L., and Marois, R. (2013). Amodal processing in human prefrontal cortex. J. Neurosci. *33*, 11573–11587.

Tark, K.-J., and Curtis, C. (2009). Persistent neural activity in the human frontal cortex when maintaining space that is off the map. Nat. Neurosci. *12*, 1463–1468.

Tombu, M., Asplund, C., Dux, P., Godwin, D., Martin, J., and Marois, R. (2011). A Unified attentional bottleneck in the human brain. Proc. Natl. Acad. Sci. USA *108*, 13426–13431.

Wandell, B., Dumoulin, S., and Brewer, A. (2007). Visual field maps in human cortex. Neuron *56*, 366–383.

Welch, R., DuttonHurt, L., and Warren, D. (1986). Contributions of audition and vision to temporal rate perception. Percept. Psychophys. *39*, 294–300.

Welch, R., and Warren, D. (1980). Immediate perceptual response to intersensory discrepancy. Psychol. Bull. *88*, 638–667.