

## RESEARCH ARTICLE

## Cortical network underlying audiovisual semantic integration and modulation of attention: An fMRI and graph-based study

Yang Xi<sup>1,2</sup>, Qi Li<sup>1\*</sup>, Ning Gao<sup>1</sup>, Siyuan He<sup>1</sup>, Xiaoyu Tang<sup>3</sup>

**1** School of Computer Science and Technology, Changchun University of Science and Technology, Changchun, Jilin, P.R. China, **2** School of Computer Science, Northeast Electric Power University, Jilin, Jilin, P.R. China, **3** School of Psychology, Liaoning Normal University, Dalian, Liaoning, P.R. China

\* liqi@cust.edu.cn



## OPEN ACCESS

**Citation:** Xi Y, Li Q, Gao N, He S, Tang X (2019) Cortical network underlying audiovisual semantic integration and modulation of attention: An fMRI and graph-based study. PLoS ONE 14(8): e0221185. <https://doi.org/10.1371/journal.pone.0221185>

**Editor:** Kewei Chen, Banner Alzheimer's Institute, UNITED STATES

**Received:** May 17, 2019

**Accepted:** July 31, 2019

**Published:** August 23, 2019

**Copyright:** © 2019 Xi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript.

**Funding:** This work was financially supported by the National Natural Science Foundation of China (grant numbers 61773076, 61806025 and 31600882), Jilin Scientific and Technological Development Program (grant numbers 20190302072GX and 20180519012JH), and Scientific Research Project of Jilin Provincial Department of Education during the 13th Five-Year Plan Period (grant number JJKH20190597KJ). The

## Abstract

Many neuroimaging and electrophysiology studies have suggested that semantic integration as a high-level cognitive process involves various cortical regions and is modulated by attention. However, the cortical network specific to semantic integration and the modulatory mechanism of attention remain unclear. Here, we designed an fMRI experiment using “bimodal stimulus” to extract information regarding the cortical activation related to the effects of semantic integration with and without attention, and then analyzed the characteristics of the cortical network and the modulating effect of attention on semantic integration. To further investigate the related cortical regions, we constructed a functional brain network for processing attended AV stimuli to evaluate the nodal properties using a graph-based method. The results of the fMRI and graph-based analyses showed that the semantic integration with attention activated the anterior temporal lobe (ATL), temporoparietal junction (TPJ), and frontoparietal cortex, with the ATL showing the highest nodal degree and efficiency; in contrast, semantic integration without attention involved a relatively small cortical network, including the posterior superior temporal gyrus (STG), Heschl's gyrus (HG), and precentral gyrus. These results indicated that semantic integration is a complex cognitive process that occurs not only in the attended condition but also in the unattended condition, and that attention could modulate the distribution of cortical networks related to semantic integration. We suggest that semantic integration with attention is a conscious process and needs a wide cortical network working together, in which the ATL plays the role of a central hub; in contrast, semantic integration without attention is a pre-attentive process and involves a relatively smaller cortical network, in which the HG may play an important role. Our study will provide valuable insights into semantic integration and will be useful for investigations on multisensory integration and attention mechanism at multiple processing stages and levels within the cortical hierarchy.

funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

The integration of information from multiple senses is a fundamental requirement for recognition of the world around us. Multisensory inputs are integrated across different stages of stimulus processing and can be modulated by attention [1–2]. Behavioral data have shown that temporally, spatially, and semantically congruent information has a facilitatory effect on performance such that bimodal stimuli are detected and discriminated faster or more accurately [3–4]. The facilitatory effect of spatial and temporal congruence (or approximate congruence) has been considered to be due to the early neural integration stages [5]. For example, it has been demonstrated in the superior colliculus of cats, termed as sensory integration [5–6]. However, this early signal-statistic-dependent sensory integration cannot account for the behavioral consequences of semantic congruency, which are termed as semantic integration [6].

Semantic integration of multisensory information is an essential cognitive process for recognizing objects and communicating effectively and has been widely studied using functional magnetic resonance imaging (fMRI) [2, 7–9]. Neuroimaging studies have shown that the integration of multisensory inputs containing semantic information involves multiple cortical regions [8–11]. For example, Ye *et al.* showed that both superior temporal regions plus the medial prefrontal cortex are involved in the integration of speech and lip movements [8]. Similarly, significant activations at the right middle and superior temporal gyri were found when localization of sound sources was semantically congruent with visual stimuli [9]. In addition, Beauchamp *et al.* found that the posterior superior temporal sulcus and middle temporal gyrus responded more strongly to semantic audiovisual stimuli than to either auditory or visual stimuli [12]. Moreover, the temporoparietal junction (TPJ), which is composed of the posterior temporal lobe and the inferior parietal lobule [13–15], and the anterior temporal lobe (ATL) [13–18] have been widely considered to play a key and special role in cross-modal semantic representation [13–18]. We speculate that the semantic information interacts within multiple cortical regions, some of which may act as hubs forming a cortical network related to the stage of semantic integration.

Activation of the cortical network may vary across different stages of multisensory integration and may be affected by attention [19]. Brains often selectively attend to some stimuli while ignoring the others, and the attended stimulus is believed to be selectively processed [20]. Selection of attention can occur in a top-down manner, which is based on the intentions and task of the observer, and can also occur in a bottom-up manner, in which attention can shift without voluntary control [1]. Neuroimaging studies have demonstrated that the brain selectively integrates the semantic information from attended objects [20] and that multisensory integration requires the objects to be fully attended [1]. In contrast, many electrophysiological studies have explored the modulation of attention on late ERP (event-related potential) components related to semantic processing and suggested that cross-modal semantic processing occurs even when the semantic information is not attended [21]. These electrophysiological studies have done much excellent work on the interaction between attention and cross-modal semantic processing, but these results cannot identify the cortical regions associated with modulation of attention, due to the low spatial resolution of the ERP method.

To date, the cortical network related to the effects of semantic integration and modulation of attention remains unclear. Many neuroimaging studies have investigated the cortical network related to semantic integration by comparing the activation related to processing semantically congruent and incongruent audiovisual stimuli [22–23]. However, it has been argued that incongruency manipulations violate natural multisensory relationships and invoke error detection processes, and their role in the characterization of natural multisensory integration

processes may be limited [2]. A more direct approach is to remove other integration effects to obtain only the effects related to semantic integration.

In the present study, we removed the effects related to early sensory integration and extracted the effect of semantic integration by utilizing “bimodal stimulus” based on a previous study [2], and then we explored the cortical network specific to the semantic integration. By comparing the activations related to the effects of semantic integration with and without attention, we discerned the modulation of attention. In addition, to further identify which region activated in semantic integration with attention is the central hub, we constructed a functional brain network for semantically congruent audiovisual stimuli presented on the attended side and analyzed the nodal properties of degree and efficiency. Our study will provide valuable insights into the stage of semantic integration, and this is important for understanding multi-sensory integration [24–26].

## Methods

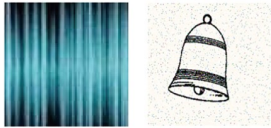
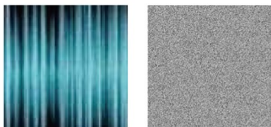

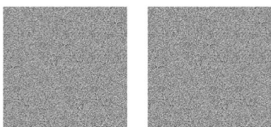
### Participants

Eighteen healthy volunteers (9 females; age range: 21–27 years; mean age: 24 years) participated in this fMRI study. All participants had normal or corrected-to-normal vision and normal hearing capability; were neurologically healthy; did not have a history or diagnosis of mental illness; did not use psychoactive medication or drugs; and did not have any permanent metal in their body. Participants were compensated with ¥100/h for the fMRI experiments. The experimental protocol was approved by the Ethics Committee of Changchun University of Science and Technology. After receiving a full explanation of the purpose and risks of the study, all participants gave written informed consent for all experiments as per the protocol approved by the institutional research review board. All procedures were carried out in accordance with the approved guidelines.

### Stimuli

A “bimodal stimulus” was constructed by using a visual or auditory noise presented simultaneously to obtain four groups: bimodal visual (Vn), bimodal auditory (An), bimodal empty (Fn), and bimodal audiovisual (AV) stimuli. Vn stimuli consisted of an auditory white noise and a visual stimulus presented simultaneously, while An stimuli consisted of a visual white noise and an auditory stimulus presented simultaneously. Fn stimuli consisted of auditory and visual white noise presented simultaneously, while AV stimuli consisted of semantically congruent auditory and visual stimuli presented simultaneously (Fig 1). Inanimate objects and animals were used as stimulus materials, since they often have distinct visual and auditory features, providing ideal stimulus sets for examining this integration process [2, 12]. Visual stimuli were gray-scale pictures of the inanimate objects (e.g., bell, guitar, car, and clock) or animals (e.g., cat, dog, cow, and frog) from the international common Snodgrass-Vanderwart white-black line graphic library and were processed using Adobe Illustrator software (Adobe Systems Inc., SAN Jose, California, USA). Finally, auditory stimuli were the sounds corresponding to the inanimate objects or animals obtained from the Internet and were processed using CoolEditPro 2.1 software (Adobe Systems Inc., SAN Jose, California, USA). Auditory noise was a white noise sound with identical length as auditory stimuli, and visual noise was a white noise image with identical size as visual stimuli. Stimulus presentation was controlled by a personal computer running Presentation 0.71 software (Neurobehavioral Systems Inc., Albany, California, USA).

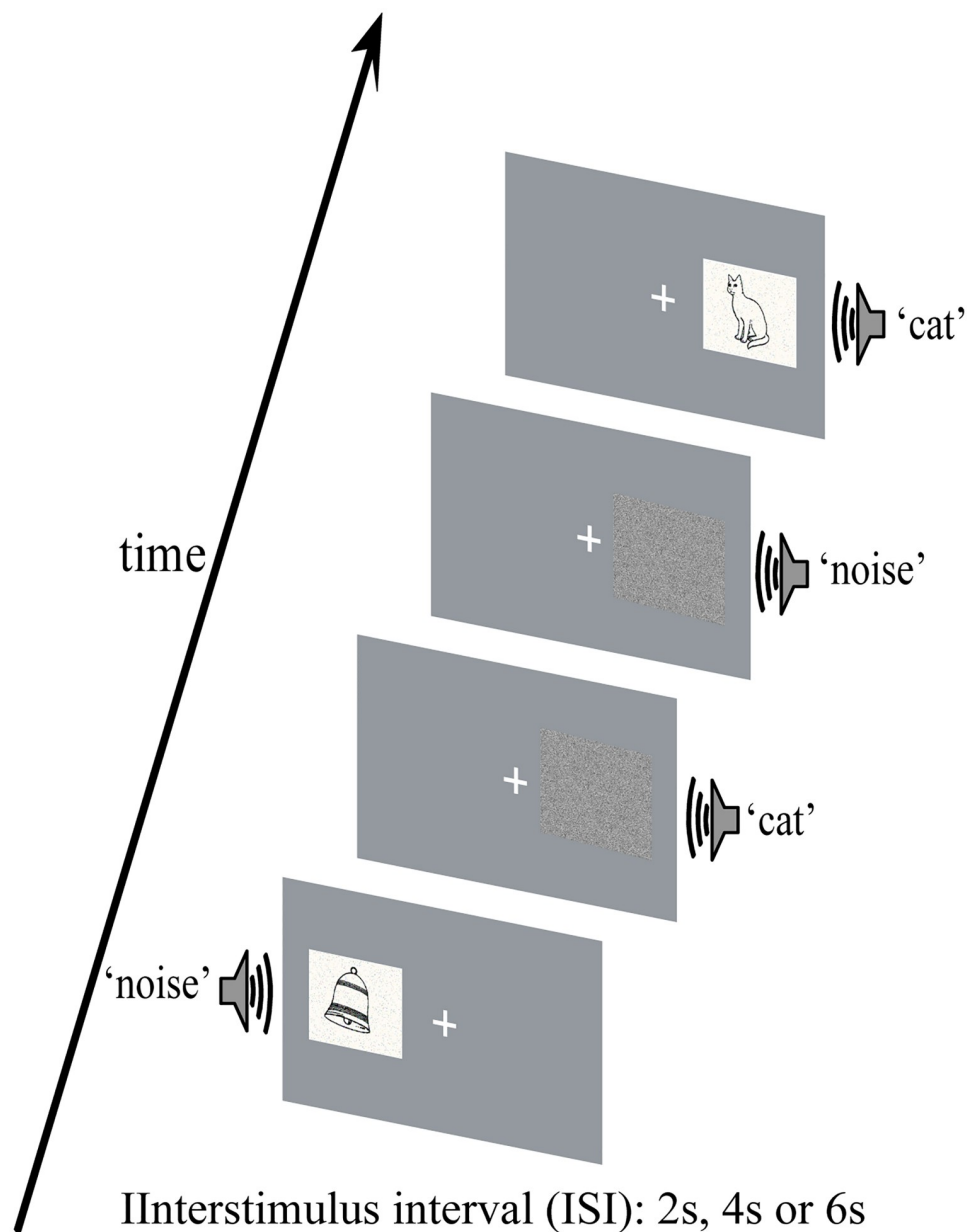
The visual stimuli and visual noise (6.0 cm × 4.8 cm, subtending a visual angle of approximately 4.3°) were projected onto a screen placed behind the head of the participants at the end

Type	Stimulus	Description
AV		Sound + Picture
An		Sound + Noise Picture
Vn		Picture + Noise Sound
Fn		Noise Picuture + Noise Sound

**Fig 1. Four types of bimodal stimuli and examples.** AV stimuli consist of semantically congruent picture and sound; An stimuli consist of a sound and a noise picture; Vn stimuli consist of a picture and a noise sound; Fn stimuli consist of noise picture and sound.

<https://doi.org/10.1371/journal.pone.0221185.g001>

of the scanner bore, visible to participants by a mirror placed within the MR head coil. All the visual stimuli were presented on the left or right side of the display at an angle of approximately  $6^\circ$  from a centrally presented fixation point located directly in front of the participants' eyes, with a duration of 300 ms. The auditory stimuli and auditory noise were presented through earphones (44 kHz sampling rate, duration 400 ms, 10 ms rise and fall periods). The hearing threshold of each participant was measured before the recordings, and the volume of the stimuli was set to 80 dB [2]. In case the participant requested, volume was increased. The interstimulus interval (ISI) varied randomly among 2 s, 4 s and 6 s. To maintain participants' attention during the discrimination task, both auditory and visual components were presented in a degraded manner [2, 27]. The visual stimuli were degraded by random noise images of the identical size. Similar to the procedure for degrading visual images, auditory stimuli were degraded by a random noise sound of identical length. The item-specific degradation level was determined based on a behavioral test before the semantic discrimination task to obtain an across-participant accuracy of 85% averaged over all items.



**Fig 2. Schematic of the experimental paradigm.** The interstimulus interval (ISI) is 2, 4 or 6 seconds.

<https://doi.org/10.1371/journal.pone.0221185.g002>

### Procedure

An event-related fMRI design was adopted. Participants continued training until the experimenter was convinced that they understood the task. Eight sessions had to be completed by each participant, and each session consisted of 40 An stimuli, 40 Vn stimuli, 40 AV stimuli, and 40 Fn stimuli, as shown in Fig 2. The frequencies of the inanimate object and animal stimuli were both 50% for each group of An, Vn and AV stimuli. There were seven types of stimuli (2 (inanimate object and animal)  $\times$  3 (An, Vn, and AV) + 1 (Fn)), which were presented with equal probability on the left and right sides of the participant according to a pseudorandom sequence.

During the experiment, participants were instructed to minimize blinking and bodily movements to avoid movement artifacts and were required to fix their eyes on a centrally presented fixation point and to attend to the stimuli on one side while ignoring the stimuli on the opposite side. The task was to press the left button when hearing or /and seeing the animal stimuli and press the right button when hearing or /and seeing the inanimate object stimuli on attended side, with responding as quickly and accurately as possible. Participants were required to attend to the left side in four of eight sessions (named as left session), and to attend to the right side in the other four sessions (named as right session). The two types of sessions were conducted in an alternating fashion. All participants were allowed to take a 5 min break between sessions.

### Acquisition of fMRI data

A 3T fMRI scanner (Siemens) at the Sino Japanese Friendship Hospital of Jilin University was used to acquire both T1-weighted anatomical images (repetition time (TR) = 8600 ms; echo time (TE) = 4 ms; field-of-view (FOV) = 192 mm; flip angle (FA) = 90°; 128 slices; voxel size = 1×1×1 mm) and T2-weighted gradient echo planar imaging sequence (TR = 2 s; TE = 30 ms; FOV = 192 mm; FA = 90°; 33 slices; voxel size = 3×3×3 mm). There were eight sessions with a total of 160 volume images per session. The high-resolution anatomical image volume was acquired at the end of the experiment.

### Data analysis

**Behavioral data analysis.** Reaction times (RTs) and hit rates (HRs) were computed separately for each type of stimulus (An, Vn and AV) and side (left and right), and analyzed using a repeated-measures analysis of variance (ANOVA) with type (An, Vn, and AV) and side (left and right) as subject factors. All statistical analysis was performed using SPSS software (version 22, IBM Inc., UAS). The  $\alpha$  level was set to 5%.

**fMRI data analysis.** The processing of each type of stimulus (An, Vn, AV and Fn) is described in Table 1. The effect of early sensory integration can be removed by comparing the audiovisual interaction in AV+Fn condition relative to those in the Vn+An condition, allowing one to extract the effect involved in the semantic integration. For all left sessions, the comparison of (AV+Fn) and (Vn+An) presented on left side reflected the effect of semantic integration with attention, while the comparison of that presented on right side reflected the effect of semantic integration without attention. The situations were similar for all right sessions. The regions associated with the modulation effect of attention were identified by comparing the semantic integration with and without attention.

The fMRI imaging data were analyzed using the SPM12 software package (Wellcome Department of Cognitive Neurology, London, UK) running under Matlab2012a (MathWorks

**Table 1. The processing of An, Vn, AV, and Fn stimuli in the brain.**

	AV	Vn	An	Fn	(AV+Fn)-(An+Vn)
Visual stimulus	✓	✓			
Auditory stimulus	✓		✓		
Visual noise			✓	✓	
Auditory noise		✓		✓	
Sensory integration	✓	✓	✓	✓	
Semantic integration	✓				✓

<https://doi.org/10.1371/journal.pone.0221185.t001>



Inc., Natick, Massachusetts, USA). Six scans at the beginning of the measurement were removed automatically from the data set. Functional data were slice time-corrected, motion-corrected, normalized into standard stereotactic space using the Montreal Neurological Institute (MNI) template, and smoothed using a 6.0-mm full-width half-maximum Gaussian kernel. To reduce motion-related artifacts, session-specific realignment parameters from preprocessing were used as first-level covariates. Statistical analysis was performed at the individual participant level by using the general linear model framework, and the blood oxygen level-dependent response was modeled as the neural activity convolved with a canonical hemodynamic response function. The contrast of (AV+Fn > An+Vn) with attention and (AV+Fn > An+Vn) without attention in all left-sessions and right-sessions were implemented. All individual functional localization data were then used for the group-level statistics. One-sample t-tests were used to construct statistical parametric maps at the group level for (AV+Fn > An+Vn) contrasts, determining the voxels in which activity differed significantly from zero, i.e., the voxels that showed significant activity in the processes of audiovisual semantic integration with and without attention.

### Graph-based analysis for the nodal properties of the functional brain network

The fMRI data corresponding to the AV stimuli presented on the attended side was entered into the CONN toolbox to construct a functional brain network [28–30]. The anatomical volumes were segmented into gray matter, white matter, and CSF areas, and the resulting masks were eroded (one voxel erosion, isotropic 2-mm voxel size) to minimize partial volume effects. The temporal time series characterizing the estimated subject motion (three-rotation and three-translation parameters, plus another six parameters representing their first-order temporal derivatives), as well as the BOLD time series within the subject-specific white matter mask (three PCA parameters) and CSF mask (three PCA parameters), were used as temporal covariates and removed from the BOLD functional data by using linear regression [28]. A network comprises nodes and edges connecting the nodes. In the present study, we defined cortical regions of interest (ROIs) from HOA112 atlas as nodes, which parcellated the brain into 112 ROIs. The mean time series for each ROI was extracted from the preprocessed images, and Pearson correlation was applied to the mean time series as the task-dependent functional connectivity between nodes. First-level (within-subjects) connectivity analysis was performed across all sessions. After computation of individual ROI-to-ROI connectivity matrices, the measure was then entered into a second-level general linear model to obtain population-level estimates and inferences. False-positive control in ROI-to-ROI analysis was implemented by using false discovery rate (FDR)-corrected p-values [28].

We entered the functional brain network into the GRETNA toolbox to perform the graph-based analysis. The graphs for ROIs and their functional connectivity were analyzed on an individual-subject basis. We selected the equal-interval sparsity threshold range (ranging from 0.05 to 0.5 with a partition interval of 0.05), and the average values of area under curve (AUC) were used for statistical analysis in order to provide a scalar that did not depend on the specific threshold selection. We computed the nodal degree and nodal efficiency, and identified the nodes with degree and efficiency values larger than the sum of the average values and the standard deviation across all nodes of the network as a hub node. In addition, we further calculated the connectivity pattern between the hub with highest nodal degree and nodal efficiency and other regions.

**Table 2. Mean response times (RTs) and hit rates (HRs).** Response times are presented in milliseconds (ms). Hit rates are represented as percentage values.

RTs (ms)		HRs (%)	
An	816.32 (99.62)	An	96.28(5.38)
Vn	735.39 (92.95)	Vn	91.92 (8.98)
AV	731.42 (88.86)	AV	98.25 (2.59)

<https://doi.org/10.1371/journal.pone.0221185.t002>

## Results

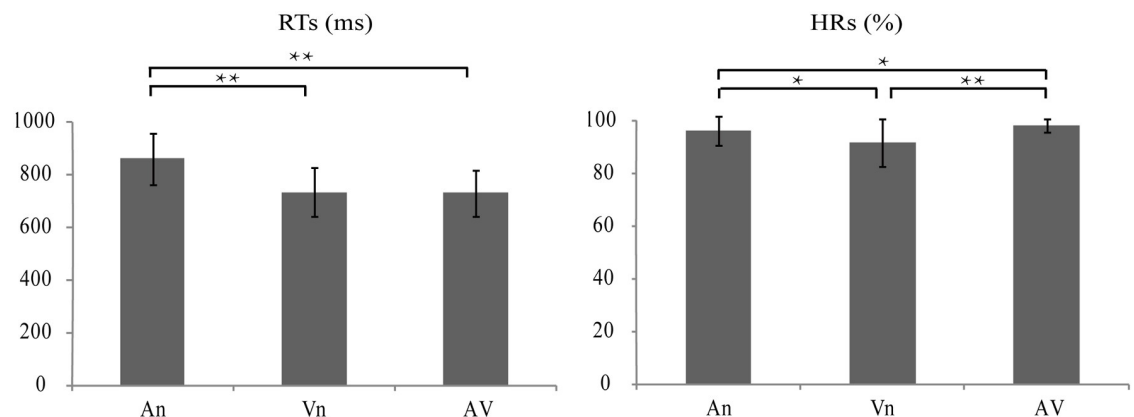
### Behavioral results

A repeated-measures ANOVA for the two factors modality (An, Vn, and AV) and hemisphere (left and right) showed no significant main effects of hemisphere in the RTs [ $F_{1,18} = 0.059$ ,  $p = 0.811$ ] or HRs [ $F_{1,18} = 1.759$ ,  $p = 0.178$ ]. Therefore, the RTs and HRs from the left and right hemispheres were combined to improve the signal-noise ratio. Mean RTs and HRs for An, Vn, and AV stimuli are shown in Table 2.

The results of RT analysis revealed main effects of modality [ $F_{2,74} = 16.774$ ,  $p < 0.0005$ ], indicating that RTs significantly differed among modalities. *Post hoc* comparisons revealed that RTs to AV stimuli were significantly faster than those to An stimuli [ $t_{37} = 18.219$ ,  $p < 0.0005$ ], and they were not significantly faster than those to Vn stimuli [ $t_{37} = 0.501$ ,  $p = 0.619$ ]. RTs to Vn stimuli were also faster than those to An stimuli [ $t_{37} = 13.943$ ,  $p < 0.0005$ ] (Fig 3). The result of HRs analysis revealed main effects of modality [ $F_{2,74} = 14.485$ ,  $p < 0.0005$ ]. *Post hoc* comparisons revealed that HRs for AV stimuli were much higher than those for An stimuli [ $t_{37} = 3.529$ ,  $p < 0.001$ ] and Vn stimuli [ $t_{37} = 5.565$ ,  $p < 0.0005$ ]. HRs for An stimuli were also significantly higher than those for Vn stimuli [ $t_{37} = 3.787$ ,  $p < 0.001$ ] (Fig 3).

### fMRI results

To identify the regions related to semantic integration with attention, we implemented the contrast of (AV+Fn) versus (An+Vn) presented on the attended side. Similarly, we obtained the effect of semantic integration without attention by implementing the contrast of (AV+Fn)



**Fig 3. Bar plots showing response times (RTs) and hit rates (HRs) for the target under the three conditions (An, Vn, and AV stimuli).** Left: RTs (across participants, mean  $\pm$  SD). Right: HRs (across participants, mean  $\pm$  SD). \*\*Significant at  $p < 0.0005$ , \*Significant at  $p < 0.05$ .

<https://doi.org/10.1371/journal.pone.0221185.g003>



**Table 3. Activated clusters related to semantic integration with attention.**

Brain region	Cluster size	Peak T	MNI coordinates (x, y, z)			Brodman area
Temporoparietal junction: Left						
Angular Gyrus	179	7.08	-48	-70	35	BA39
Supramarginal Gyrus		5.21	-54	-64	29	BA 40
Middle Temporal Gyrus		5.07	-39	-58	23	BA 21
Superior Temporal Gyrus		4.30	-41	-59	20	BA 22
Temporoparietal junction: Right						
Angular Gyrus	128	5.75	51	-70	32	BA 39
Supramarginal Gyrus		4.51	51	-55	29	BA 40
Supramarginal Gyrus		4.49	57	-55	23	BA 40
Middle Temporal Gyrus		4.76	53	-63	26	BA 21
Superior Temporal Gyrus		4.53	56	-61	26	BA 22
Anterior temporal lobe: Left						
Middle Temporal Gyrus	51	4.78	-48	8	-37	BA 21
Middle Temporal Gyrus		4.61	-42	14	-40	BA 21
Middle Temporal Gyrus		4.39	-39	5	-40	BA 21
Middle Temporal Gyrus	37	4.56	-57	-7	-19	BA 21
Middle Temporal Gyrus		4.21	-60	-16	-13	BA 21
Anterior temporal lobe: Right						
Middle Temporal Gyrus	101	5.10	48	8	-34	BA 21
Middle Temporal Gyrus		4.77	54	-4	-25	BA 21
Middle Temporal Gyrus		4.48	54	5	-25	BA 21
Superior Temporal Gyrus		4.51	45	11	-31	BA 22
Frontal cortex						
Medial Superior Frontal Gyrus	367	7.06	9	53	44	BA 8
Medial Frontal Gyrus		5.38	15	44	53	BA 8/9
Medial Frontal Gyrus		5.36	-12	38	56	BA 8/9
Superior Frontal Gyrus		4.45	-12	48	41	BA 4/6/8
Medial Superior Frontal Gyrus	132	6.37	-12	47	-16	BA 8
Medial Orbitofrontal cortex		4.63	0	35	-13	BA 10/11
Inferior Frontal Gyrus		4.70	3	44	-16	BA 11
Parietal cortex						
Precuneus	35	6.29	-24	-85	38	BA 7
Precuneus		4.91	-24	-82	41	BA 7
Posterior Cingulate Gyrus		4.67	-6	-52	26	BA 31
Lateral Occipital gyrus	43	5.35	30	-82	38	BA 19
Superior parietal lobule		4.49	30	-76	50	BA 5/7
Precuneus		4.08	21	-82	44	BA 7
Posterior Cingulate Gyrus	144	5.28	9	-49	26	BA 31

$P_{\text{uncorr}} < 0.0005$ ; cluster voxels  $\geq 30$  voxels.

<https://doi.org/10.1371/journal.pone.0221185.t003>

versus (An+Vn) presented on the unattended side. Tables 3 and 4 show the significantly activated clusters related to semantic integration with and without attention.

The effect of semantic integration with attention was located in the bilateral temporoparietal junction (TPJ), anterior temporal lobe (ATL), frontal lobe, and parietal lobe cortices. Specifically, the posterior middle temporal gyrus, posterior superior temporal gyrus, angular gyrus,

**Table 4. Activated clusters related to semantic integration with attention.**

Brain region	Cluster size	Peak T	MNI coordinates (x, y, z)			Brodmann area
			x	y	z	
Precentral Gyrus	90	5.25	-36	-19	59	BA 4
Precentral Gyrus		4.90	-30	-25	65	BA 4
Precentral Gyrus		4.02	-24	-25	59	BA 4
Supramarginal Gyrus	87	4.81	-51	-28	14	BA 40
Heschl's Gyrus		4.58	-45	-25	8	BA 41
Superior Temporal Gyrus		4.42	-48	-25	8	BA 22
Heschl's Gyrus	31	4.36	48	-22	11	BA 41
Superior Temporal Gyrus		4.01	48	-21	8	BA 22

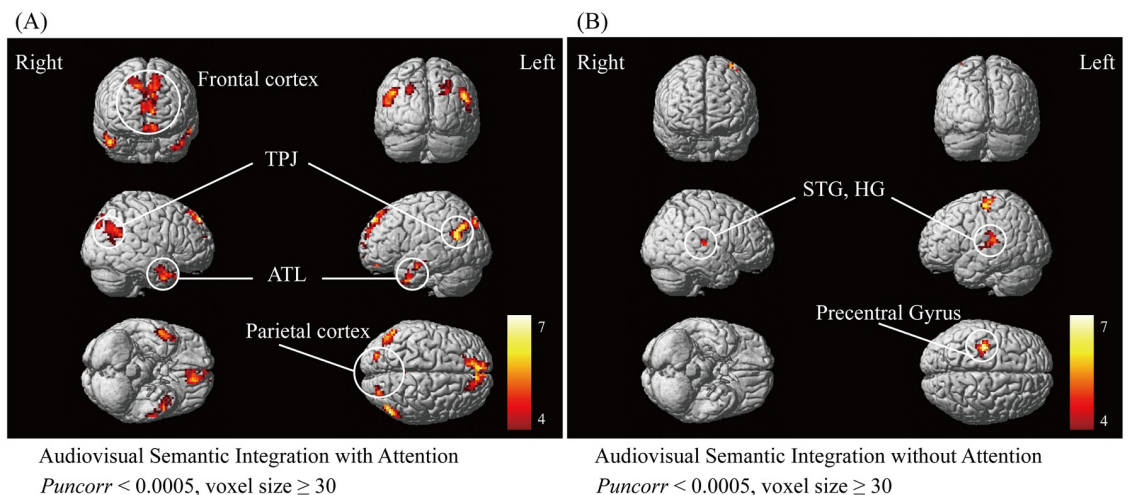
$P_{uncorr} < 0.0005$ ; cluster voxels  $\geq 30$  voxels.

<https://doi.org/10.1371/journal.pone.0221185.t004>

supramarginal gyrus in the TPJ, anterior middle temporal gyrus, anterior superior temporal gyrus in the ATL, medial frontal gyrus, superior frontal gyrus, inferior frontal gyrus and orbitofrontal cortex in the frontal lobe cortex, precuneus, and the posterior cingulate gyrus and superior parietal lobule in the parietal lobe cortex were strongly activated, as shown in Fig 4(A). The activation results indicated that the cortical network related to semantic integration with attention is widely distributed in the prefrontal, parietal and temporal lobes. In contrast, the significant activation corresponding to semantic integration without attention involved a relatively small cortical network, mainly including the posterior superior temporal gyrus (STG), Heschl's gyrus (HG), and precentral gyrus, as shown in Fig 4(B).

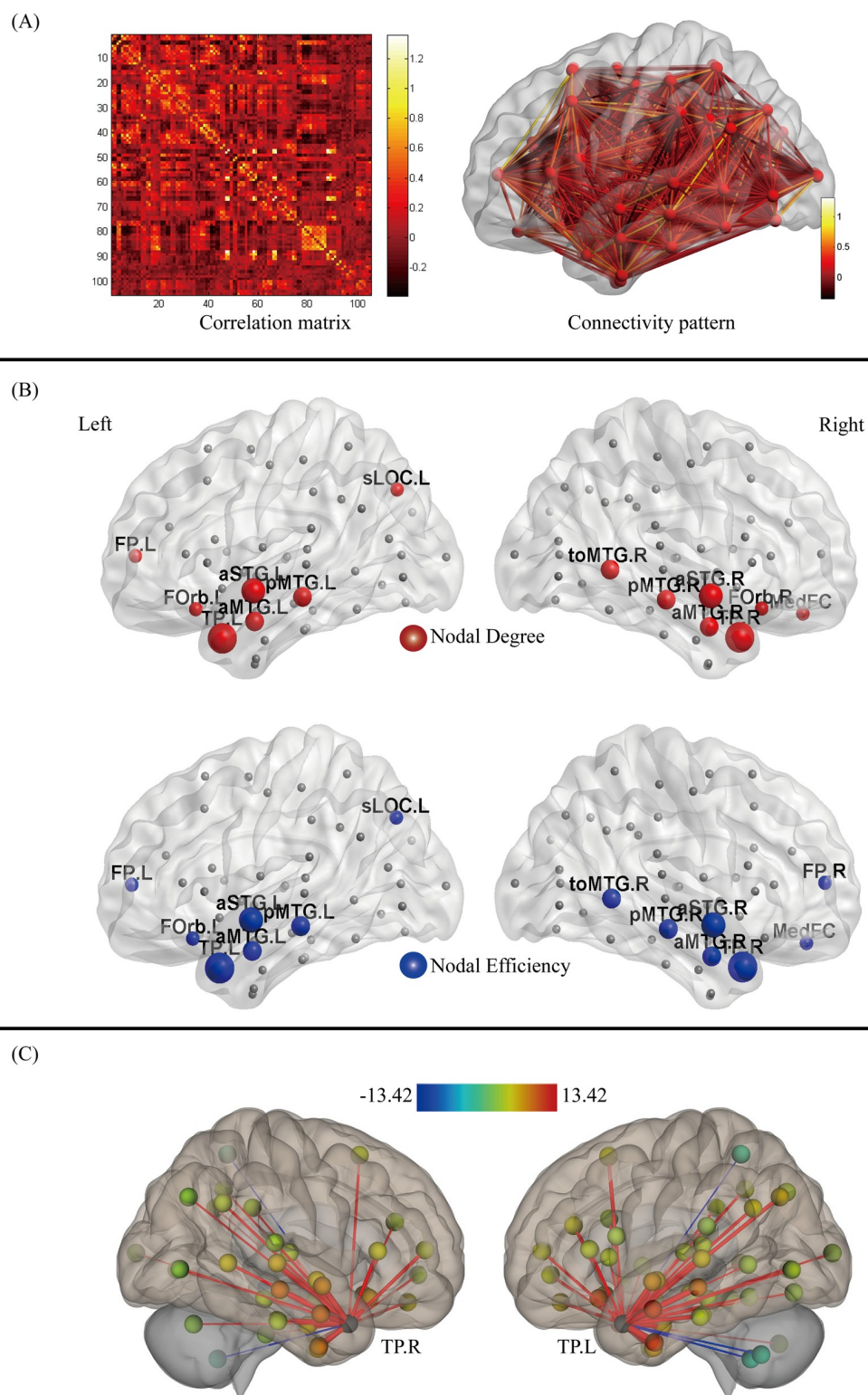
### Nodal degree and efficiency

By constructing the functional brain network, we obtained the correlation matrix and the functional connectivity pattern, as shown in Fig 5(A). Then we computed the nodal degree and efficiency and obtained the hubs, values of which were shown in Table 5.



**Fig 4. (A) The fMRI activation observed in contrast of (AV+Fn) > (An+Vn) presented on the attended side, reflecting the cortical network of semantic integration with attention. (B) The fMRI activation observed in contrast of (AV+Fn) > (An+Vn) presented on the unattended side, reflecting the cortical network of semantic integration without attention.  $P_{uncorr} < 0.0005$ , cluster voxels  $\geq 30$  voxels. TPJ, temporoparietal junction; ATL, anterior temporal lobe; STG, superior temporal gyrus; HG, Heschl's gyru. The color bar indicates T values.**

<https://doi.org/10.1371/journal.pone.0221185.g004>



**Fig 5. (A) The correlation matrix and the functional connectivity pattern in the brain for processing AV stimuli attended.** The color bar indicates the strength of connectivity. **(B) The distribution of hubs with high degree and nodal efficiency.** The node that degree or efficiency value is larger than the sum of the average value and the standard deviation across all nodes of the network is defined as a hub node. The red spheres are the hubs with high nodal degree, and the blue spheres are the hubs with high nodal efficiency. The size of the spheres indicates the relative size of the nodal degree or efficiency. **(C) The strength of functional connectivity between bilateral temporal poles (TP.L and TP.R) with other regions in temporal lobes.** The color bar indicates the strength of connectivity.

<https://doi.org/10.1371/journal.pone.0221185.g005>

Table 5. The information of hubs with high nodal degree and nodal efficiency.

Node	Description	Degree	Node	Description	Efficiency
TP.L	left temporal pole	17.607	TP.L	left temporal pole	0.2909
TP.R	right temporal pole	17.144	TP.R	right temporal pole	0.2878
aSTG.L	left anterior superior temporal gyrus	16.641	aSTG.L	left anterior superior temporal gyrus	0.2851
aSTG.R	right anterior superior temporal gyrus	16.354	aSTG.R	right anterior superior temporal gyrus	0.2827
pMTG.L	left posterior middle temporal gyrus	16.146	pMTG.L	left posterior middle temporal gyrus	0.2822
pMTG.R	right posterior middle temporal gyrus	15.964	pMTG.R	right posterior middle temporal gyrus	0.2812
toMTG.R	right temporo-occipital middle temporal gyrus	15.880	toMTG.R	right temporo-occipital middle temporal gyrus	0.2806
aMTG.L	left anterior middle temporal gyrus	15.660	aMTG.L	left anterior middle temporal gyrus	0.2794
aMTG.R	right anterior middle temporal gyrus	15.651	aMTG.R	right anterior middle temporal gyrus	0.2785
FOrb.L	left frontal orbital cortex	15.481	sLOC.L	left lateral occipital cortex	0.2775
sLOC.L	left lateral occipital cortex	15.352	FOrb.L	left frontal orbital cortex	0.2773
MedFC	frontal medial cortex	15.087	FP.L	left frontal pole	0.2746
FP.L	left frontal pole	14.964	MedFC	frontal medial cortex	0.2737
FOrb.R	right frontal orbital cortex	14.799	FP.R	right frontal pole	0.2736

<https://doi.org/10.1371/journal.pone.0221185.t005>

The distribution of hubs with high nodal degree and nodal efficiency is shown in Fig 5(B), and the values of nodal degree and efficiency of hubs are shown in Table 5. The result showed that the fourteen top degree nodes were mainly concentrated in the temporal and frontal lobes, including the left temporal pole (TP.L), right temporal pole (TP.R), left anterior superior temporal gyrus (aSTG.L), right anterior superior temporal gyrus (aSTG.R), left posterior middle temporal gyrus (pMTG.L), right posterior middle temporal gyrus (pMTG.R), right temporo-occipital middle temporal gyrus (toMTG.R), left anterior middle temporal gyrus (aMTG.L), the right anterior middle temporal gyrus (aMTG.R), left frontal pole (FP.L), left frontal orbital cortex (FOrb.L), right frontal orbital cortex (FOrb.R), frontal medial cortex (MedFC), and left superior lateral occipital cortex (sLOC.L). Interestingly, as shown in Fig 5(B), the hubs with high efficiency were almost identical to those with high degree, including left temporal pole (TP.L), right temporal pole (TP.R), left anterior superior temporal gyrus (aSTG.L), right anterior superior temporal gyrus (aSTG.R), left posterior middle temporal gyrus (pMTG.L), right posterior middle temporal gyrus (pMTG.R), right temporo-occipital middle temporal gyrus (toMTG.R), left anterior middle temporal gyrus (aMTG.L), right anterior middle temporal gyrus (aMTG.R), left superior lateral occipital cortex (sLOC.L), left frontal orbital cortex (FOrb.L), left frontal pole (FP.L), frontal medial cortex (MedFC), right frontal pole (FP.R). In addition, the values of nodal degree and efficiency of these hubs decreased gradually from the anterior temporal lobe to the posterior temporal lobe. Most importantly, the distribution showed that the two nodes with the highest degree and efficiency were both TP.L and TP.R. Further analysis showed that the connectivity strength between TP in the left and right hemispheres was the highest, followed by the connectivity between TP and other regions in the temporal lobe, including aSTG.R, aSTG.L, aMTG.R, aMTG.L, pMTG.R, pMTG.L, aITG.R, and aITG.L, and the connectivity strength between the TP and other regions in the temporal lobe decreased gradually from the anterior to the posterior temporal lobe, as shown in Fig 5(C).

## Discussion

The present study aimed to identify the cortical networks underlying audiovisual semantic integration and explore the modulatory mechanism of attention. Behavioral evidence for audiovisual semantic integration was obtained from HRs and RTs. RTs to AV stimuli were

significantly faster than those to An stimuli, while HRs to AV stimuli were significantly higher than those to both An and Vn stimuli. The behavioral facilitation of RTs and HRs might indicate that congruent semantic information within AV stimuli facilitated cross-modal integration and further improved the performance of the semantic discrimination task [31].

### The cortical network related to audiovisual semantic integration with attention

By implementing the  $(AV+Fn) > (An+Vn)$  contrast presented on the attended side, we obtained the cortical activation corresponding to the effect of semantic integration with attention (Fig 4(A)).

Two important regions of the TPJ and ATL were strongly activated during semantic integration with attention. These areas were widely discussed in multiple studies as association cortices of semantic representation [32–33]. The TPJ mainly contains the posterior temporal lobe and inferior parietal lobule. Compared to the unisensory auditory and visual inputs, AV stimulation induced stronger activation in the posterior temporal lobe, indicating that this region was involved in the semantic representation of cross-modal objects [34]. Similarly, Beauchamp *et al.* reported that the posterior temporal lobe showed an enhanced response when semantic auditory and visual object were presented together, relative to presentation in a single modality [12]. In addition, some studies demonstrated that the inferior parietal lobule, consisting of the angular gyrus and supramarginal gyrus, is an important region for cross-modal semantic representation [14, 35–36]. Geschwind and co-workers argued that the inferior parietal lobule is ideally connected as a cross-modal central region to code the semantic information contained in words [37]. In contrast, there has been an accumulation of functional neuroimaging [14, 38–39] studies of the ATL, which have found that the ATL is activated for a range of semantic tasks, irrespective of the modality of input (e.g., words, pictures, sounds, etc.) [40]. Moreover, temporary interference with ATL activity can produce semantic impairments across a range of tasks with words and pictures [41–42]. To further support these findings, conceptual impairments have been found in semantic dementia patients with ATL atrophy [43]. Here, TPJ and ATL were significantly activated, providing direct evidence that they are both involved in semantic integration. However, the role of these two regions as central hubs in cross-modal semantic representation has always been a matter of debate [14–18, 33, 44]. In the present study, the analysis of the nodal properties of the functional brain network showed that the two nodes with the highest degree and the highest nodal efficiency are the bilateral temporal pole (TP.L and TP.R) located in the ATL (Fig 5(B)). The high nodal efficiency reflects that the nodes connected to hubs are more tightly clustered together, and the high degree reflects that these nodes had an increased influence on general network processing [45]. The bilateral temporal pole regions showed the highest degree and nodal efficiency, indicating that these regions might contribute more to the information interaction efficiently. Moreover, the values of nodal degree and nodal efficiency decrease gradually from the anterior to posterior temporal lobe (Fig 5(B)), and the pattern of functional connectivity between bilateral TP and other regions within the temporal lobe showed that the connectivity strength gradually increased from posterior, middle, to the anterior areas, with the connectivity between the left and right temporal poles being the strongest (Fig 5(C)). Supporting our findings, neuroscience studies have indicated that convergence of sensory information in the temporal lobe is a graded process that occurs along both its longitudinal and lateral axes and culminates in the most rostral limits [46–47]. Thus, we suggested that both TPJ and ATL are important sites for semantic integration with attention, but ATL may serve as a central hub role in the cortical network.



Another significant activation related to semantic integration in the attention condition was found in the frontal and parietal cortices. The frontal cortex was demonstrated to belong to wide cortical networks representing cross-modal semantic associations [48–50], which receives afferent connections from multiple association areas, and integrates cross-modal information in support of behavior [48–50]. For example, an fMRI study showed that the frontal cortex was activated more strongly during semantically congruent compared to incongruent AV stimulation [49]. Naumer *et al.* found that multisensory stimuli of sounds and images with semantic information was integrated in the frontal cortex [50], suggesting that the frontal cortex is responsible for semantic information integration. In addition, FP, FOrb and MedFC in frontal cortex are the hubs with high nodal degree and nodal efficiency (Fig 5(B)), reflecting that frontal cortex contributes more to effective semantic integration with attention. Beyond the frontal cortex, the parietal cortex was also found to be similarly activated. Some studies have suggested that multisensory information would be integrated and represented in the frontal lobe via the parietal lobe [51–52], which is generally thought of as a higher-order association area and was observed to be activated in semantically congruent audiovisual integration [53]. These studies supported our finding of the activation in the frontal and parietal lobes during semantic integration, reflecting that these areas are both engaged in cross-modal semantic representation.

### The cortical network related to audiovisual semantic integration without attention

By implementing the  $(AV+Fn) > (An+Vn)$  contrast presented on the unattended side, we observed that the STG, HG and precentral gyrus were strongly activated (Fig 4(B)). Some ERP studies have suggested that cross-modal semantic processing occurs even when the semantic information is not actively attended [54]. Our present result supports this contention and further identifies the cortical regions responsible for semantic integration without attention.

The specificity of HG for semantic integration under the unattended condition was an interesting finding. This primary auditory cortex was activated in semantic integration with attention but not activated without attention, indicating that it may play an important role in unattended semantic integration. Neurophysiological studies in human and nonhuman primates have found early integration in the primary auditory cortex [55], and suggested that multisensory integration processes in the primary sensory cortices are governed by tight temporal [56] and spatial constraints [57]. However, in the present study, we removed the effect of early integration induced by a congruent temporospatial relationship through the experimental design with bimodal stimuli, and focused on the effect of semantic integration. Thus, our result showed that higher semantic integration occurred in the primary auditory cortex, suggesting that the multisensory integration in the primary auditory cortex is also related to a congruent semantic relationship.

Multisensory interactions in the primary auditory cortex can be mediated by several types of functional neural architectures, including feedforward thalamocortical, direct connections between sensory areas and feedback from higher-order association areas [58–59]. We suggested that one possibility is that the semantic information from the auditory and visual modalities was input into the HG from subcortical architectures, and then was integrated. Alternatively, the auditory and visual signals were combined in the association areas, such as the posterior STG here, and the subsequent outcome affected the response amplification in HG by means of feedback projections [60].

Our results indicated that HG is an important region for cross-modal semantic representation in the absence of attention. Another significant activation was noted in the precentral



gyrus, which was suggested to be associated with decision-making of movement [52]. We speculated that this may reflect the mechanism of cognitive control, which inhibited the participants from responding to the stimuli on the unattended side by pressing a button. However, further studies are needed to confirm and elucidate these findings.

### The modulating effect of attention on semantic integration

By comparing the results of semantic integration with and without attention, we observed that in comparison with the unattended condition, when attention was allocated and directed to the stimuli, the integration sites were widely distributed in the temporal, frontal, and parietal lobes, indicating that attention strongly affected semantic integration processing [61–62]. First, attention modulated the response in the frontoparietal cortices, which were strongly activated in semantic integration with attention but were barely activated in integration without attention. This result supported the theory that frontoparietal cortex is a source of top-down attention [61], adjusting the neuron responses to favor information that is currently relevant for behavior [62]. This selective representation may serve as a source of bias, prioritizing the processing of task-relevant information across the brain [63]. We speculated that the allocation of attention resource enabled the neurons in the frontoparietal cortex to focus on processing the task-related stimuli presented on the attended side. Secondly, attention modulated the cross-modal semantic representation in the temporal cortex, which enhanced the response in ATL but inhibited the response in HG. The ATL as a central hub was specific to semantic integration with attention, which was not activated in the unattended condition, and the HG was specific to semantic integration without attention, which was not observed to be activated in the attended condition. The posterior STG was activated both in attended and unattended conditions, indicating that this region may be a generally site for semantic integration. These differences might indicate that ATL is responsible for conscious semantic integration that needs allocation of attention. In other words, attention is a prerequisite for semantic integration in ATL. In contrast, the semantic integration in HG did not require attention to be allocated, which might reflect an unconscious or a pre-attentive processing [26, 64]. We suggest that the activation patterns of cortical networks indicated that semantic integration is a complex cognitive process, which may contain multiple integration processes occurring in different cortical regions and are modulated by attention. However, further experimentation is required to clarify our speculation.

### Conclusion

In the present study, we explored the cortical networks underlying semantic integration and the modulation of attention by using fMRI and graph-based methods. The results showed that the cortical network related to semantic integration with attention was distributed in the frontal, parietal, and temporal lobes, while the cortical activation related to semantic integration without attention was located in the posterior STG, HG, and precentral gyrus. The different responses in cortical regions reflected that semantic integration could occur both in attended and unattended conditions, and attention can facilitate semantic integration to enhance cross-modal semantic representation [61–62]. We suggest that semantic integration with attention is a conscious process and needs a wide cortical network working together, in which ATL plays the role of a central hub. Semantic integration without attention is a pre-attentive process and involves a relatively small cortical network, in which the HG may play an important role. However, one limitation of our study is that our sample size was relatively small. It is important to validate our findings by replicating our analyses in a larger sample of subjects. Our present

study would be useful to investigate multisensory integration and attention at multiple processing stages and levels within the cortical hierarchy.

## Acknowledgments

The authors would like to thank all individuals who participated in the initial experiments from which raw data were collected.

## Author Contributions

**Formal analysis:** Ning Gao, Siyuan He, Xiaoyu Tang.

**Funding acquisition:** Qi Li.

**Investigation:** Yang Xi, Qi Li.

**Writing – original draft:** Yang Xi, Qi Li.

## References

1. Talsma D, Senkowski D, Soto-Faraco S, Woldorff MG. The multifaceted interplay between attention and multisensory integration. *Trends Cogn Sci*. 2010; 14(9):400–10. <https://doi.org/10.1016/j.tics.2010.06.008> PMID: 20675182
2. Werner S, Noppeney U. Distinct Functional Contributions of Primary Sensory and Association Areas to Audiovisual Integration in Object Categorization. *Journal of Neuroscience*. 2010; 30(7):2662–75. <https://doi.org/10.1523/JNEUROSCI.5091-09.2010> PMID: 20164350
3. Van der Burg E, Olivers CN, Bronkhorst AW, Theeuwes J. Pip and pop: nonspatial auditory signals improve spatial visual search. *J Exp Psychol Hum Percept Perform*. 2008; 34(5):1053–65. <https://doi.org/10.1037/0096-1523.34.5.1053> PMID: 18823194
4. Van der Burg E, Olivers CN, Bronkhorst AW, Theeuwes J. Poke and pop: tactile-visual synchrony increases visual saliency. *Neurosci Lett*. 2009; 450(1):60–4. PMID: 19013216
5. Meredith MA, Stein BE. Spatial determinants of multisensory integration in cat superior colliculus neurons. *J Neurophysiol*. 1996; 75(5):1843–57. <https://doi.org/10.1152/jn.1996.75.5.1843> PMID: 8734584
6. Meyer GF, Greenlee M, Wuergler S. Interactions between Auditory and Visual Semantic Stimulus Classes: Evidence for Common Processing Networks for Speech and Body Actions. *J Cognitive Neurosci*. 2011; 23(9):2291–308. <https://doi.org/10.1162/jocn.2010.21593> PMID: 20954938
7. Werner S, Noppeney U. Superadditive Responses in Superior Temporal Sulcus Predict Audiovisual Benefits in Object Categorization. *Cerebral Cortex*. 2010; 20(8):1829–42. <https://doi.org/10.1093/cercor/bhp248> PMID: 19923200
8. Ye Z, Russeler J, Gerth I, Munte TF. Audiovisual Speech Integration in the Superior Temporal Region Is Dysfunctional in Dyslexia. *Neuroscience*. 2017; 356:1–10. <https://doi.org/10.1016/j.neuroscience.2017.05.017> PMID: 28527953
9. Plank T, Rosengarth K, Song W, Ellermeier W, Greenlee MW. Neural correlates of audio-visual object recognition: Effects of implicit spatial congruency. *Hum Brain Mapp*. 2012; 33(4):797–811. <https://doi.org/10.1002/hbm.21254> PMID: 21425399
10. Sadaghiani S, Maier JX, Noppeney U. Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. *J Neurosci*. 2009; 29(20):6490–9. <https://doi.org/10.1523/JNEUROSCI.5437-08.2009> PMID: 19458220
11. Noesselt T, Tyll S, Boehler CN, Budinger E, Heinze HJ, Driver J. Sound-induced enhancement of low-intensity vision: multisensory influences on human sensory-specific cortices and thalamic bodies relate to perceptual enhancement of visual detection sensitivity. *Journal of Neuroscience the Official Journal of the Society for Neuroscience*. 2010; 30(41):13609–13623. <https://doi.org/10.1523/jneurosci.4524-09.2010> PMID: 20943902
12. Beauchamp MS, Lee KE, Argall BD, Martin A. Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*. 2004; 41(5):809–23. [https://doi.org/10.1016/S0896-6273\(04\)00070-4](https://doi.org/10.1016/S0896-6273(04)00070-4) PMID: 15003179
13. Lambon Ralph MA, Jefferies E, Patterson K, Rogers TT. The neural and computational bases of semantic cognition. *Nat Rev Neurosci*. 2016; 18:42–55. <https://doi.org/10.1038/nrn.2016.150> PMID: 27881854

14. Schwartz MF, Kimberg DY, Walker GM, Brecher A, Faseyitan OK, Dell GS, et al. Neuroanatomical dissociation for taxonomic and thematic knowledge in the human brain. *Proc Natl Acad Sci U S A*. 2011; 108(20):8520–4. <https://doi.org/10.1073/pnas.1014935108> PMID: 21540329
15. Mirman D, Landrigan JF, Britt AE. Taxonomic and thematic semantic systems. *Psychol Bull*. 2017; 143(5):499–520. <https://doi.org/10.1037/bul0000092> PMID: 28333494
16. Jackson RL, Hoffman P, Pobric G, Lambon Ralph MA. The Nature and Neural Correlates of Semantic Association versus Conceptual Similarity. *Cereb Cortex*. 2015; 25(11):4319–33. <https://doi.org/10.1093/cercor/bhv003> PMID: 25636912
17. Chiou R, Lambon Ralph MA. Unveiling the dynamic interplay between the hub- and spoke-components of the brain's semantic system and its impact on human behavior. *NeuroImage*. 2019; 199(1):114–126. <https://doi.org/10.1016/j.neuroimage.2019.05.059> PMID: 31132452
18. Adolfo MG, Sebastian M, Kathya T, Enrique G, Eduar H, Edinson M, et al. How meaning unfolds in neural time: Embodied reactivations can precede multimodal semantic effects during language processing. *NeuroImage*. 2019; 197(15): 439–449. <https://doi.org/10.1016/j.neuroimage.2019.05.002> PMID: 31059796
19. Talsma D, Woldorff MG. Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *Journal of Cognitive Neuroscience*. 2005; 17(7): 1098–1114. <https://doi.org/10.1162/0898929054475172> PMID: 16102239
20. Li Y, Long J, Huang B, Yu T, Wu W, Li P, et al. Selective Audiovisual Semantic Integration Enabled by Feature-Selective Attention. *Sci Rep*. 2016; 6:18914. <https://doi.org/10.1038/srep18914> PMID: 26759193
21. Talsma D, Doty TJ, Woldorff MG. Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cereb Cortex*. 2007; 17(3):679–90. <https://doi.org/10.1093/cercor/bhk016> PMID: 16707740
22. Naghavi HR, Eriksson J, Larsson A, Nyberg L. Cortical regions underlying successful encoding of semantically congruent and incongruent associations between common auditory and visual objects. *Neurosci Lett*. 2011; 505(2):191–5. <https://doi.org/10.1016/j.neulet.2011.10.022> PMID: 22024505
23. Li Y, Wang G, Long J, Yu Z, Huang B, Li X, et al. Reproducibility and discriminability of brain patterns of semantic categories enhanced by congruent audiovisual stimuli. *PLoS One*. 2011; 6(6): e20801. <https://doi.org/10.1371/journal.pone.0020801> PMID: 21750692
24. Horn A, Ostwald D, Reiser M, Blankenburg F. The structural-functional connectome and the default mode network of the human brain. *Neuroimage*. 2014; 102(1):142–51. <https://doi.org/10.1016/j.neuroimage.2013.09.069>
25. Zuo XN, Ehmke R, Mennes M, Imperati D, Castellanos FX, Sporns O, et al. Network centrality in the human functional connectome. *Cereb Cortex*. 2012; 22(8):1862–75. <https://doi.org/10.1093/cercor/bhr269> PMID: 21968567
26. Zhou FA, Davis G. Unconscious priming of task sets: the role of spatial attention. *Atten Percept Psychophys*. 2012; 74(1):105–14. <https://doi.org/10.3758/s13414-011-0221-8> PMID: 21972047
27. Stevenson RA, James TW. Audiovisual integration in human superior temporal sulcus: inverse effectiveness and the neural processing of speech and object recognition. *Neuroimage*. 2009; 44(3): 1210–1223. <https://doi.org/10.1016/j.neuroimage.2008.09.034> PMID: 18973818
28. Whitfield-Gabrieli S, Nieto-Castanon A. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect*. 2012; 2(3):125–41. <https://doi.org/10.1089/brain.2012.0073> PMID: 22642651
29. Katherine W, Jeff A, Jennifer T, Todd F, Anthony B, Christof K. Characterization of Functional Brain Connectivity towards Optimization of Music Selection for Therapy—A fMRI study. *International Journal of Neuroscience*. 2019; 129(9):882–889. <https://doi.org/10.1080/00207454.2019.1581189> PMID: 30744538
30. Mireia H, Noelia V, Albert C, Anna M, César Ávila. Brain networks involved in accented speech processing. *Brain and Language*. 2019; 194: 12–22. <https://doi.org/10.1016/j.bandl.2019.03.003> PMID: 30959385
31. Li Q, Yu H, Li X, Sun H, Yang J, Li C. The informativity of sound modulates crossmodal facilitation of visual discrimination: a fMRI study. *Neuroreport*. 2017; 28(2):63–8. <https://doi.org/10.1097/WNR.0000000000000719> PMID: 27926629
32. Mahon BZ, Caramazza A. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*. 2008; 102(1):59–70. <https://doi.org/10.1016/j.jphysparis.2008.03.004>

33. Lambon Ralph MA. Neurocognitive insights on conceptual knowledge and its breakdown. *Philos Trans R Soc Lond B Biol Sci.* 2014; 369(1634):20120392. <https://doi.org/10.1098/rstb.2012.0392> PMID: [24324236](https://pubmed.ncbi.nlm.nih.gov/24324236/)
34. Stevenson R, Geoghegan M, James T. Superadditive BOLD activation in superior temporal sulcus with threshold non-speech objects. *Exp. Brain Res.* 2007; 179:85–95. <https://doi.org/10.1007/s00221-006-0770-6> PMID: [17109108](https://pubmed.ncbi.nlm.nih.gov/17109108/)
35. Schwartz MF, Kimberg DY, Walker GM, Brecher A, Faseyitan OK, Dell GS, et al. Neuroanatomical dissociation for taxonomic and thematic knowledge in the human brain. *P Natl Acad Sci USA.* 2011; 108(20):8520–4. <https://doi.org/10.1073/pnas.1014935108> PMID: [21540329](https://pubmed.ncbi.nlm.nih.gov/21540329/)
36. Mirman D, Landrigan JF, Britt AE. Taxonomic and thematic semantic systems. *Psychol Bull.* 2017; 143(5):499–520. <https://doi.org/10.1037/bul0000092> PMID: [28333494](https://pubmed.ncbi.nlm.nih.gov/28333494/)
37. Binder JR, Desai RH. The neurobiology of semantic memory. *Trends Cogn Sci.* 2011; 15(11):527–36. <https://doi.org/10.1016/j.tics.2011.10.001> PMID: [22001867](https://pubmed.ncbi.nlm.nih.gov/22001867/)
38. Lambon Ralph MA, Sage K, Jones RW, Mayberry EJ. Coherent concepts are computed in the anterior temporal lobes. *Proc Natl Acad Sci U S A.* 2010; 107(6):2717–22. <https://doi.org/10.1073/pnas.0907307107> PMID: [20133780](https://pubmed.ncbi.nlm.nih.gov/20133780/)
39. Jackson RL, Hoffman P, Pobric G, Lambon Ralph MA. The Nature and Neural Correlates of Semantic Association versus Conceptual Similarity. *Cereb Cortex.* 2015; 25(11):4319–33. <https://doi.org/10.1093/cercor/bhv003> PMID: [25636912](https://pubmed.ncbi.nlm.nih.gov/25636912/)
40. Patterson K, Nestor PJ, Rogers TT. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci.* 2007; 8(12):976–87. <https://doi.org/10.1038/nrn2277> PMID: [18026167](https://pubmed.ncbi.nlm.nih.gov/18026167/)
41. Pobric G, Jefferies E, Lambon Ralph MA. Category-specific versus category-general semantic impairment induced by transcranial magnetic stimulation. *Curr Biol.* 2010; 20(10):964–8. <https://doi.org/10.1016/j.cub.2010.03.070> PMID: [20451381](https://pubmed.ncbi.nlm.nih.gov/20451381/)
42. Pobric G, Jefferies E, Ralph MA. Amodal semantic representations depend on both anterior temporal lobes: evidence from repetitive transcranial magnetic stimulation. *Neuropsychologia.* 2010; 48(5):1336–42. <https://doi.org/10.1016/j.neuropsychologia.2009.12.036> PMID: [20038436](https://pubmed.ncbi.nlm.nih.gov/20038436/)
43. Rogers TT, Patterson K, Graham K. Colour knowledge in semantic dementia: it is not all black and white. *Neuropsychologia.* 2007; 45(14):3285–98. <https://doi.org/10.1016/j.neuropsychologia.2007.06.020> PMID: [17804024](https://pubmed.ncbi.nlm.nih.gov/17804024/)
44. Binney RJ, Parker GJ, Lambon Ralph MA. Convergent connectivity and graded specialization in the rostral human temporal lobe as revealed by diffusion-weighted imaging probabilistic tractography. *J Cogn Neurosci.* 2012; 24(10):1998–2014. [https://doi.org/10.1162/jocn\\_a\\_00263](https://doi.org/10.1162/jocn_a_00263) PMID: [22721379](https://pubmed.ncbi.nlm.nih.gov/22721379/)
45. Burdette JH, Laurienti PJ, Espeland MA, Morgan AT, Telesford QK, Vechlekar CD, et al. Using network science to evaluate exercise-associated brain changes in older adults. *Frontiers in Aging Neuroscience.* 2010; 2(23):23. <https://doi.org/10.3389/fnagi.2010.00023> PMID: [20589103](https://pubmed.ncbi.nlm.nih.gov/20589103/)
46. Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci.* 2009; 12(6):718–24. <https://doi.org/10.1038/nn.2331> PMID: [19471271](https://pubmed.ncbi.nlm.nih.gov/19471271/)
47. Deco G, Rolls ET. A neurodynamical cortical model of visual attention and invariant object recognition. *Vision Res.* 2004; 44(6):621–42. <https://doi.org/10.1016/j.visres.2003.09.037> PMID: [14693189](https://pubmed.ncbi.nlm.nih.gov/14693189/)
48. Zhang Y, Wang K, Yue C, Gao S, Huang P, Wang T, et al. Prefrontal sensitivity to changes in language form and semantic content during speech production. *Brain and Language.* 2019; 194:23–34. <https://doi.org/10.1016/j.bandl.2019.04.001> PMID: [30991263](https://pubmed.ncbi.nlm.nih.gov/30991263/)
49. Laurienti PJ, Wallace MT, Maldjian JA, Susi CM, Stein BE, Burdette JH. Cross-modal sensory processing in the anterior cingulate and medial prefrontal cortices. *Hum Brain Mapp.* 2003; 19(4):213–23. <https://doi.org/10.1002/hbm.10112> PMID: [12874776](https://pubmed.ncbi.nlm.nih.gov/12874776/)
50. Naumer MJ, Doehrmann O, Muller NG, Muckli L, Kaiser J, Hein G. Cortical plasticity of audio-visual object representations. *Cereb Cortex.* 2009; 19(7):1641–53. <https://doi.org/10.1093/cercor/bhn200> PMID: [19015373](https://pubmed.ncbi.nlm.nih.gov/19015373/)
51. Okada K, Hickok G. Two cortical mechanisms support the integration of visual and auditory speech: a hypothesis and preliminary data. *Neurosci Lett.* 2009; 452(3):219–23. <https://doi.org/10.1016/j.neulet.2009.01.060> PMID: [19348727](https://pubmed.ncbi.nlm.nih.gov/19348727/)
52. Huang J, Reinders A, Wang Y, Xu T, Zeng YW, Li K, et al. Neural correlates of audiovisual sensory integration. *Neuropsychology.* 2018; 32(3):329–36. <https://doi.org/10.1037/neu0000393> PMID: [29620404](https://pubmed.ncbi.nlm.nih.gov/29620404/)
53. Calvert GA, Campbell R, Brammer MJ. Evidence from functional magnetic resonance imaging of cross-modal binding in the human heteromodal cortex. *Curr Biol.* 2000; 10(11):649–57. [https://doi.org/10.1016/s0960-9822\(00\)00513-3](https://doi.org/10.1016/s0960-9822(00)00513-3) PMID: [10837246](https://pubmed.ncbi.nlm.nih.gov/10837246/)

54. Relander K, RaMa P, Kujala T. Word semantics is processed even without attentional effort. *Journal of Cognitive Neuroscience*. 2009; 21(8):1511–1522. <https://doi.org/10.1162/jocn.2009.21127> PMID: [18823236](https://pubmed.ncbi.nlm.nih.gov/18823236/)
55. Laurienti PJ, Burdette JH, Wallace MT, Yen YF, Field AS, Stein Be. Deactivation of sensory-specific cortex by cross-modal stimuli. *J Cogn Neurosci*. 2002; 14:420–429. <https://doi.org/10.1162/089892902317361930> PMID: [11970801](https://pubmed.ncbi.nlm.nih.gov/11970801/)
56. Kayser C, Petkov CI, Logothetis NK. Visual modulation of neurons in auditory cortex. *Cereb Cortex*. 2008; 18:1560–1574. <https://doi.org/10.1093/cercor/bhm187> PMID: [18180245](https://pubmed.ncbi.nlm.nih.gov/18180245/)
57. Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron*. 2007; 53:279–292. <https://doi.org/10.1016/j.neuron.2006.12.011> PMID: [17224408](https://pubmed.ncbi.nlm.nih.gov/17224408/)
58. Schroeder CE, Smiley J, Fu KG, McGinnis T, O'Connell MN, Hackett TA. Anatomical mechanisms and functional implications of multisensory convergence in early cortical processing. *Int J Psychophysiol*. 2003; 50:5–17. [https://doi.org/10.1016/s0167-8760\(03\)00120-x](https://doi.org/10.1016/s0167-8760(03)00120-x) PMID: [14511832](https://pubmed.ncbi.nlm.nih.gov/14511832/)
59. Driver J, Noesselt T. Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron*. 2008; 57:11–23. <https://doi.org/10.1016/j.neuron.2007.12.013> PMID: [18184561](https://pubmed.ncbi.nlm.nih.gov/18184561/)
60. Calvert GA, Brammer MJ, Bullmore ET, Campbell R, Iversen SD, David AS. Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport*. 1999; 10(12):2619–23. <https://doi.org/10.1097/00001756-199908200-00033> PMID: [10574380](https://pubmed.ncbi.nlm.nih.gov/10574380/)
61. Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci*. 2002; 3(3):201–15. <https://doi.org/10.1038/nrm755> PMID: [11994752](https://pubmed.ncbi.nlm.nih.gov/11994752/)
62. Duncan J. The structure of cognition: attentional episodes in mind and brain. *Neuron*. 2013; 80(1):35–50. <https://doi.org/10.1016/j.neuron.2013.09.015> PMID: [24094101](https://pubmed.ncbi.nlm.nih.gov/24094101/)
63. Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci*. 2001; 24:167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167> PMID: [11283309](https://pubmed.ncbi.nlm.nih.gov/11283309/)
64. Weibel S, Giersch A, Dehaene S, Huron C. Unconscious task set priming with phonological and semantic tasks. *Conscious Cogn*. 2013; 22(2):517–27. <https://doi.org/10.1016/j.concog.2013.02.010> PMID: [23558083](https://pubmed.ncbi.nlm.nih.gov/23558083/)