Research Article　　　　　　　　　　　　　　　　　**Open Access**

The Neural Correlates of Auditory Lexicosemantic Processing in the Chinese Language: An fMRI Study

Hengshuang LIU[1] and SH Annabel CHEN[2*]

[1]National Key Research Centre for Linguistics and Applied Linguistics, Bilingual Cognition and Development Lab, Guangdong University of Foreign Studies, Guangzhou, China

[2]Psychology, School of Social Sciences, Centre for Research and Development in Learning, Lee Kong Chian School of Medicine (LKCMedicine), Nanyang Technological University, Singapore

*Correspondence to:

S. H. Annabel Chen, PhD
Professor of Psychology, School of Social Sciences, Nanyang Technological University
48 Nanyang Avenue, Singapore 639818
Tel: (65) 6316-8836
Fax: (65) 6795-5797
E-mail: aannabelchen@ntu.edu.sg

Abstract

Everyday social communication emphasizes auditory semantic processing. To date, most studies of auditory semantic processing are mainly based on languages with alphabetic writing systems like English, where character languages such as Chinese are largely under-represented. Thus, the current study attempted to investigate the auditory semantic neuro-network for the Chinese language using fMRI. Twenty-three native Mandarin Chinese speakers were scanned while performing an auditory semantic-tone task. Results indicate that the Chinese auditory lexicosemantic network may involve the bilateral posterior superior temporal lobes, the left middle frontal gyrus, the left ventral inferior frontal gyrus, the anterior superior temporal cortex, the left middle temporal gyrus, and the left occipito-temporal cortex. Within this Chinese auditory lexicosemantic network, co-activity was observed between the right posterior superior temporal cortex and frontal areas. This verifies the significance of the fronto-temporal connectivity in the Chinese lexicosemantic network specifically in the auditory domain. Overall, the current network is consistent with the classical language networks based on alphabetic writing systems, but with features specific to Chinese. While the left inferior parietal lobule underlying sublexical-level phonemic assembly may not be involved in Chinese auditory lexicosemantic processing, the left middle frontal gyrus is likely to be recruited for the sound-to-form mapping at the whole character level. Taken together, the current study deepens our understanding of how the linguistic and neurobiological representations interact during Chinese auditory lexicosemantic processing.

Keywords

Chinese, Character languages, Auditory lexicosemantic network, Spoken words, Functional connectivity

Introduction

One of the core processes in oral communication is understanding the meanings of words expressed. It is particularly interesting to identify how the brain processes the auditory semantic information in the Chinese language since radicals carrying semantic cues in Chinese characters apparently do not aid in this process. To our knowledge, this issue has only been investigated in four studies published to date [1-4]. Based on a review of these studies, Chinese auditory lexicosemantic processing involves the bilateral posterior superior temporal lobes, the left middle frontal gyrus, the left ventral inferior frontal gyrus, the left anterior

superior temporal cortex, the left middle temporal gyrus, and the left occipito-temporal cortex. The current study will revisit these four past studies and further elaborate on the Chinese auditory lexicosemantic network using a neuroimaging study.

The earliest neuroimaging investigation among these four studies [3] found that the bilateral occipital-temporal cortices (BA37) and the bilateral middle temporal gyri (BA21) had greater activity when making auditory lexical decisions on real disyllabic Chinese words (e.g., '领村'/tai4 yang2/ sun) than on pseudowords (e.g.,'太阳'/ling3 cun1/). It is surprising to capture the lexical effect in the bilateral occipital-temporal visual cortices, as no visual word forms were physically presented. The authors thereafter argued that this finding would reflect the automatic activation of Chinese visual characters when corresponding phonological representations were activated. In addition to the bilateral visual cortices, the bilateral middle temporal gyri were also seen in the words > pseudowords contrast. This possibly indicated the greater lexicosemantic representations in real words than in pseudowords. Despite the recruitment of the middle temporal gyrus, it was still difficult to estimate to what extent this lexical decision task involved semantic-related activation. As participants were not explicitly required to make decisions based on the semantic quality of the heard words, the pseudowords can be rejected simply due to the unfamiliar phonology without the necessity to exploit semantic-level processing [5]. Essentially, Xiao et al. [3] also claimed that the focus of their study was on phonological forms instead of semantic representations of two-character Chinese words.

More recently, consistent findings are observed in two studies (Liu et al., using similar task paradigms, where auditory meaning relatedness judgment compared to tone baseline evoked greater activations in the bilateral posterior temporal lobes, the left occipito-temporal cortex [1-4], and the left ventral frontal lobe. Both studies recognized the significance of the left pars triangularis (ventral inferior frontal gyrus) in Chinese auditory lexicosemantic processing. It has also been interpreted in these two studies that the recruitment of the left occipito-temporal cortex was driven by the interaction of orthographic and phonological representations during spoken Chinese word recognition. However, the statistical power of the region-of-interest (ROI) results in Zou et al.'s study [4] could be stronger if the ROIs were defined from separate studies rather than from the same study.

In addition to regional activation patterns, interregional connectivity network underpinning Chinese auditory lexicosemantic processing was examined in Wu et al.'s study [2] by using multivariate independent component analysis (ICA). While the regional activation results observed in the occipito-temporal cortex (right lingual gyrus, BA18) and the left ventral frontal lobe (BA45/47) were consistent with Liu et al.'s [1] and Zou et al. [4], several strongly-connected networks were identified within the extensive fronto-temporal cortex when contrasting Chinese auditory semantic dangerousness judgment (e.g., '手枪'/shou3 qiang1/ gun) with a resting baseline. This highlights the significance of the fronto-temporal co-activation in Chinese auditory lexicosemantic

processing. However, it was also recognized by Wu et al. [2] that the networks were extracted in a relatively broad manner, without being further separated into more precise sub-networks such as interlinks between two paired regions. Nevertheless, this study broadens our understanding of the complex neuro-mechanism underlying Chinese auditory lexicosemantic processing, which is not only mediated by isolated brain regions but also dependent on the interactions of several areas in a parallel distributed hierarchy.

Apart from Wu et al.'s study [2], most other studies examining the effective/functional connectivity underlying Chinese lexicosemantic processing utilize visually- rather than aurally-presented tasks. For example, Fan et al. [6, 7] used dynamic causal modelling (DCM) and found that semantic relatedness judgment of visual Chinese characters appeared to be modulated by several interregional connections. The bidirectional connectivity between the left ventral inferior frontal gyrus and the left middle temporal gyrus suggests that the ventral inferior frontal gyrus may assist to retrieve semantic knowledge stored in the left middle temporal gyrus, while the connection from the left fusiform gyrus to the left middle temporal gyrus signals the bottom-up influences of orthography on semantic representations. It thus has implications for the left middle temporal gyrus to serve as an interface to integrate top-down controlled and bottom-up automatic processes.

In addition to studies investigating task-based effective/functional connectivity, the anatomical connectivity underlying the Chinese semantic system has also been examined using diffusion tensor imaging [8]. In this study, reduced fractional anisotropy and greater lesion percentage in the left inferior fronto-occipital fasciculus correlated with more severe semantic deficits in Chinese patients, irrespective of factors such as the input modality (visual vs. auditory), output modality (verbal vs. non-verbal), patients' overall cognitive state, whole lesion volume, type of brain damage, and gray matter involvement. This indicates that the Chinese semantic system is likely to be regulated by the inferior fronto-occipital fasciculus connecting the posterior lateral temporal cortex to the frontal cortex and underscores the significance of this fiber tract for semantic processing in Chinese.

As mentioned earlier, there are limitations in these four studies that have implications for the motivation of the current investigation. For example, Xiao et al.'s study [3] utilized a lexical decision task that may not intensely engage semantic network; the ROIs in Zou et al.'s study [4] were not a priori but were defined using conjunction results in their own study; limited studies investigated the ROI-to-ROI functional connectivity underlying Chinese lexicosemantic processing in the auditory modality. Given this, an empirical investigation is needed to evaluate the Chinese auditory lexicosemantic activations by employing a task that explicitly requires semantic-based responses, and at the same time to examine functional connectivity between a priori ROIs. To achieve this, the current study utilized an fMRI data set, where native Mandarin Chinese participants were required to judge whether each aurally presented Chinese word denoted

a living creature bigger than a cat. A priori regions of interest (ROIs) were created based on the past four studies [1-4], and functional connectivity between these ROIs was computed. It is hypothesized that activations and functional connectivity will be observed within the classical language networks that include the bilateral posterior superior temporal lobes, the left middle frontal gyrus, the left ventral inferior frontal gyrus, the left anterior superior temporal cortex, and the left middle temporal gyrus, as well as regions related to visual word forms, the left occipito-temporal cortex within the fusiform gyrus.

Materials and Methods

Participants

An archival neuroimaging data set of 26 native speakers of Mandarin Chinese was employed in this study. The study was conducted in Taiwan and all participants used traditional Chinese characters in their daily reading and writing. None of them had a history of neurological diseases or psychiatric disorders, and the participants had received 17.6 years of education (SD = 2.2 years, range = 13-23 years) on average. All participants gave informed consent approved by the Institutional Review Board at National Taiwan University Hospital before the experiment.

Three male participants were excluded due to poor image quality. The remaining 23 right-handed participants (12 females; mean age = 25.7 years, SD = 4.2 years, range = 19-35 years; mean handedness score = 87.8, SD = 17.6, range = 42.8-100) were included in data analysis.

Procedure

A blocked design auditory semantic-tone task, which was adapted from Binder et al. [9], was applied in the present study. In the semantic condition, the participants were instructed to judge whether a heard disyllabic Chinese noun described a living creature bigger than a cat. Two- instead of one-character nouns were adopted to lessen the incidence of homophones and then to avoid the potential semantic confusion arising from irrelevant homophones. As a baseline, the tone condition required the participants to determine whether there were two high tones (750 Hz) amongst a sequence of three to seven tones (either 250Hz or 750Hz).

After performing a ten-trial practice outside the scanner, participants were administered the formal task in one run while being scanned. The run consisted of ten 30s blocks that alternated between the semantic condition and the tone condition, forming five cycles each comprising one semantic and one tone block (Figure 1). Each block had ten 3s trials. In each trial, after an auditory stimulus was played, participants were required to respond with a button press using their right hand before the next trial started.

Image acquisition

The current experiment applied functional magnetic resonance imaging (fMRI). This technology is non-invasive and replicable with high spatial resolution down to millimeters and fair temporal resolution within a few seconds. Given that neural activity is accompanied by cerebral blood flow, fMRI measures neural activity by detecting variations of the oxygen level in blood. That is, when blood flow to a brain region increases, the neural activity of this region is presumed to increase accordingly.

Participants were scanned in a 3 Tesla Siemens Trio MRI scanner with a 12-channel head coil at National Taiwan University Hospital. The echo planar imaging (EPI) sequence was used to obtain functional images: repetition time (TR) = 2000 ms, echo time (TE) = 24 ms, flip angle (FA) = 90 °, field of view (FOV) = 240 mm, matrix size = 64 × 64, voxel size = 3.8 mm x 3.8 mm x 3.8 mm, slice thickness = 3.8 mm, and 34 axial slices aligned to the anterior-posterior commissure plane with a total of 150 images per task. A T2-weighted image was acquired with TR = 5920 ms, TE = 102 ms, FA = 150°, FOV = 250mm, matrix size = 256 × 256, voxel size = 1 mm x 1 mm x 3.9 mm, slice thickness = 3.9 mm, and 34 axial slices. A high resolution T1-weighted 3D Magnetization-Prepared RApid Gradient-Echo (MPRAGE) whole-brain scan was also acquired using TR = 1380 ms, TE = 2.6 ms, FA = 15°, FOV = 250 mm, matrix size = 256 × 256, and voxel size = 1mm x 1mm x 1mm, slice thickness = 1 mm.

Image analysis

The neuroimaging data were subjected to a three-stage pre-processing protocol [10] implemented in Statistical Parametric Mapping 8 (SPM8; Department of Cognitive Neurology, London, UK). The first stage was conducted for each participant, where all structural and functional images were reoriented to an origin at the anterior commissure, followed by a slice timing correction of the functional images to the middle slice [11], a realignment of the functional images to the first volume, a co-registration of the T1 MPRAGE image to the mean functional image, and a segmentation of the T1 MPRAGE image into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). This first stage generated a GM template and a WM template for each participant. The GM and WM templates from all participants were superimposed together in the second stage, to generate a group template and a flow field for each participant. The group template and individual flow field were utilized for the normalization and smoothing of each participant in the last stage. Individual functional images and GM images were normalized to Montreal Neurological Institute (MNI) space and smoothed with a Gaussian kernel of 8 mm full-width at half-maximum (FWHM). While the first pre-processing stage went through conventional pre-processing, the second and third stages underwent Diffeomorphic Anatomical Registration using Exponentiated Lie algebra (DARTEL) pre-processing instead for improved normalization [12-13]. After the three stages of pre-processing, whole-brain

Figure 1: The condition sequence in the task run.
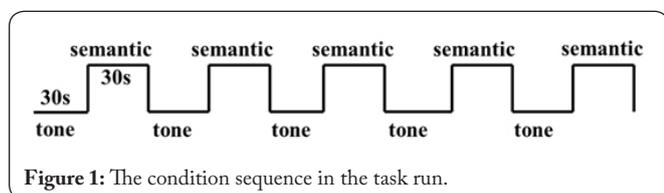
activations in the semantic > tone contrast were obtained for each participant using the first-level analysis in SPM8. For second-level analysis, group activation across all participants was computed for the semantic > tone contrast using a one-sample t-test.

In addition to the whole-brain analyses, ROIs were also defined for the examination of the possible functional connections between them. Based on the past four studies [1-4], the left middle frontal gyrus, the left ventral inferior frontal gyrus, the left anterior superior temporal gyrus, the left middle temporal gyrus, the bilateral posterior superior temporal gyri, and the left occipito-temporal cortices were selected as a priori ROIs. The centre coordinates of these ROIs were identified from a meta-analysis conducted by the first author [14]. Using the MarsBaR toolbox [15], the ROIs of the left posterior superior temporal gyrus, the right posterior superior temporal gyrus, the left anterior superior temporal gyrus, and the left occipito-temporal cortex were defined as 8mm spheres, centred at (-58, -14, -2), (64, -6, -6), (-50, 14, -16), and (-56, -54, -14), respectively. The ROIs of the left middle frontal gyrus, the left ventral inferior frontal gyrus, and the left middle temporal gyrus were defined as 8mm spheres, centred at (-48, 24, 24), (-42, 36, -4), and (-60, -42, -2), respectively.

Functional connectivity between every pair of these predefined ROIs was calculated in the semantic > tone contrast by CONN 17b (https://www.nitrc.org/projects/conn) implemented in SPM8. ROI-to-ROI functional connectivity indicates the level of linear association of the blood oxygen level-dependent (BOLD) time series between each pair of ROIs [16-17]. When distributed brain ROIs display strongly correlated patterns of neural activity change, it is taken as evidence that these ROIs are functionally connected [18]. After 'setup' and 'denoising' phases, the first (individual) level analysis generates files for all the possible participant/condition/ROI combinations. In the second (group) level analysis, by specifying Fisher z-transformed correlation coefficient values for between-condition and between-ROI contrasts, the corresponding participant/condition/ROI files yielded in the first-level would be extracted together for group-level correlation analyses accordingly. It is believed that the ROI-to-ROI functional connectivity represents the best approach to directly reveal brain connectivity [19-20]. Here, our study utilized the most updated version (17b) of the CONN toolbox at the time when this analysis was conducted, where the component-based noise correction (CompCor) method is implemented to improve the analysis sensitivity, selectivity, and inter-scan reliability [21].

Results

Two out of the 23 participants had behavioural data missing. Another two participants had only their practice performance available, with above 70% accuracy for both of them. Of the remaining 19 participants, 11 participants had technical issues with the performance data recorded. They had 10-50% trials with undefined 's' responses rather than the expected 'yes' or 'no' responses. This is presumably caused by the task programming where the 'yes' and 'no' were not

exclusively set as the only two allowable response types, which permitted some actual responses to be erroneously overwritten by scanning pulses. Given this, the trials with 's' responses were taken as missing trials rather than incorrect trials and then were excluded from the actual computation of accuracy and reaction time. In addition to the 's' trials, missing responses were found in all the 19 participants that ranged from 6% to 56% of the trials. However, the recorded trials (excluding the 's' trials) regardless of their correctness were verified to be distributed throughout the entire run for each of the 19 participants, indicating that these 19 participants were most likely engaged in the whole run without being noticeably distracted until the end of the run. The main purpose of collecting the behavioural data was to make sure that the participants were engrossed in the task during the whole scan session, which was not intended to be used for data analysis. Thus, we conducted some data checks below to ensure activation maps of the participants with complete behavioural data recording did not differ from those participants whose data were partially recorded.

Ten participants had more than 50% trials recorded and an accuracy rate of higher than 70% in the recorded trials, thus they were combined into one group, while the remaining 13 participants were taken as another group for having more problematic behavioural data. An independent sample t-test was conducted between the imaging data of these two groups, and no significant difference in functional activations was found between groups with a threshold of $p < .001$ uncorrected, minimum of 20 voxels. This suggests that the 13 participants with problematic response recordings are likely to be equally involved in the semantic-tone task as with the ten participants with better behavioural data quality, and thus their imaging data could be a valid representation of the targeted semantic aspect. In addition, the behavioural data was not needed for the neuroimaging analysis. Therefore, all 23 participants were included in imaging analyses to obtain a reasonable sample size with adequate statistical power.

As illustrated in table 1 and figure 2, whole-brain analysis of the 23 participants showed that the semantic > tone condition induced greater cortical activations in the left superior/medial frontal gyrus (BA8/10), the left supplementary motor area (BA6), the right cerebellum (Crus I/II), the left middle occipital gyrus (BA19), the left superior parietal lobule (BA7), the right superior temporal gyrus (BA38/48), the left hippocampus, the left pars triangularis (BA48), the left middle frontal gyrus (BA44/9), the left inferior temporal gyrus
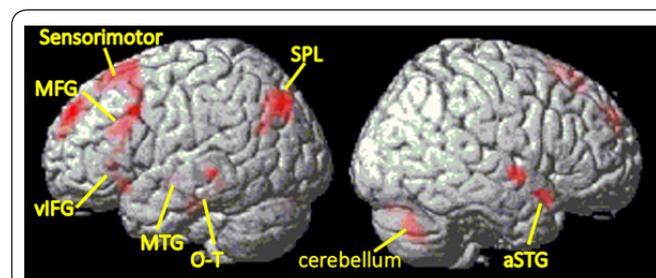

Figure 2: The whole-brain activation maps, p(FWE) < 0.05, cluster size ≥ 10 voxels.

Table 1: Peak coordinates within the significant clusters of the whole brain activation in the semantic > tone contrast.

Region	BA	x	y	z	Cluster size	T
Left superior frontal gyrus	9, 10	-15	53	34	59	9.41
Left superior/medial frontal gyri, Left supplementary motor area	8, 6	-8	27	61	94	7.94
Right cerebellum (Crus I/II)	--	42	-68	-34	32	8.39
Left middle occipital gyrus	19	-30	-76	38	94	8.02
Left superior parietal lobule	7	-27	-72	46	32	7.83
Right superior temporal gyrus	38	53	11	-15	11	7.94
Left hippocampus	--	-27	-11	-11	28	7.81
Right superior temporal gyrus	48	61	-4	-4	16	7.68
Left ventral inferior frontal gyrus (pars triangularis)	48	-46	23	27	60	7.54
Left middle frontal gyrus	44, 9	-49	15	38	27	7.09
Left inferior temporal gyrus	20	-38	-23	-23	22	7.43
Left temporal fusiform gyrus	20	-34	-30	-19	18	6.44
Left middle temporal gyrus	21	-49	-34	-8	16	6.95
Left ventral inferior frontal gyrus (pars orbicularis)	47	-34	19	-15	14	6.84

x, y, z coordinates: MNI space; BA: Brodmann Area; Cluster criterion: p(FWE) < 0.05, cluster size ≥ 10 voxels.

(BA20), the left temporal fusiform gyrus (BA20), the left middle temporal gyrus (BA21), and the left pars orbicularis (BA47), with a threshold of p(FWE) < 0.05 and cluster size ≥ 10 voxels. The relatively conservative threshold of p(FWE) < 0.05 could diminish false-positive rates, while a cluster extent threshold of 10 voxels may avoid false alarms of overly small and unreliable effects. This threshold has recently been demonstrated to be valid and has the potential to produce replicable and dependable results [22, 23]. Corresponding brain regions and Brodmann areas (BAs) of resulted coordinates were identified using respective templates in the MRIcron toolbox (https://www.nitrc.org/projects/mricron) .

Significant functional connectivity in the semantic > tone contrast was found between the left ventral inferior frontal gyrus and the right posterior superior temporal gyrus, t(22) = 3.18, p = 0.004, between the left middle frontal gyrus and the right posterior superior temporal gyrus, t(22) = 2.81, p = 0.010, and also between the left ventral inferior frontal gyrus and the left middle frontal gyrus, t(22) = 2.32, p = 0.030, under a threshold of p < 0.05 (Table 2, Figure 3).

Discussion

The current study aims to identify the Chinese auditory lexicosemantic network using a semantic-tone task. Based on the whole-brain activation and functional connectivity results, Chinese auditory lexicosemantic network is likely to involve
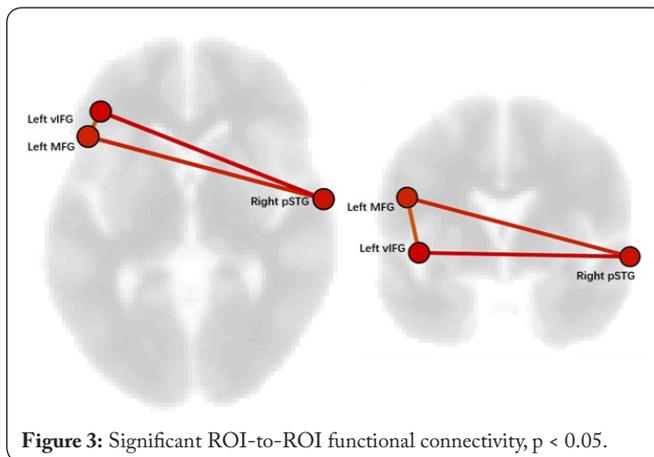

Figure 3: Significant ROI-to-ROI functional connectivity, p < 0.05.

Table 2: Significant ROI-to-ROI functional connectivity in the semantic > tone contrast.

ROI-to-ROI Connectivity	t(22)	p
Left ventral inferior frontal gyrus - right posterior superior temporal gyrus	3.18	0.004
Right posterior superior temporal gyrus - left middle frontal gyrus	2.81	0.01
Left ventral inferior frontal gyrus - left middle frontal gyrus	2.32	0.03

Threshold: p < 0.05.

the bilateral posterior superior temporal lobes, the left middle frontal gyrus, the left ventral inferior frontal gyrus, the left anterior superior temporal cortex, the left middle temporal gyrus, and the left occipito-temporal cortex.

Whole-brain activation results

As expected, lexicosemantic > tone processing was seen to engage several regions, including the left occipito-temporal cortex (BA19/20), the left middle temporal gyrus (BA21), the right anterior superior temporal cortex (BA38/48), the left ventral inferior frontal gyrus (BA47), and the left middle frontal gyrus (BA44/9). The bilateral posterior superior temporal cortices that were originally hypothesized were not observed, while unexpected activations were noted in the left superior parietal lobule (BA7), the left sensorimotor area (BA6/8/9/10), and the right cerebellum (Crus I/II).

In the current study, activations in the left occipito-temporal cortex were observed in the left temporal fusiform gyrus (BA20), the left inferior temporal gyrus (BA20), and the left middle occipital gyrus (BA19). Past studies have suggested that BA20 is more associated with orthographic representation at the higher-order linguistic level, rather than the non-linguistic visual imagery of the heard object, as visual feature extract is usually regulated by the primary visual cortex, BA17, an area posterior to BA20 [24-30]. Hence, the observed activation in the left BA20 is believed to reflect the orthographic analysis involved in spoken Chinese lexicosemantic processing rather than task-induced visual imagery. This suggests the possibility of visualization processes when listening to Chinese words, given that the Chinese visual form has greater clarity, reliability, and precision than that of

the spoken form in discriminating homophones and semantic decoding. Consistent with past findings [1-4], this suggests that orthographic information may be processed even upon hearing a word without the necessity of seeing it.

In addition to the left occipito-temporal cortex, temporal lobe activations were also observed. It included the left middle temporal gyrus, a repository site of lexicosemantic representations [31-32], as well as the right anterior superior temporal cortex (BA38/48), an area underlying basic syntactic construction such as category registration [29, 32-34]. It is notable that the anterior superior temporal lobe observed in the current study was located in the right rather than the left hemisphere. Likewise, the right anterior superior temporal lobe has also been reported in Wu et al.'s study [2], which used a semantic dangerousness task that required word categorization judgment (dangerous or not). This is not surprising based on previous findings [32-33], where word category information was believed to be stored in the anterior superior temporal gyrus bilaterally.

In the frontal lobe, activations were shown in the left ventral inferior frontal gyrus, an area supporting controlled retrieval of stored semantic representations, as well as the left middle frontal gyrus, a region thought to be specialized for lexical coordination in the Chinese language [1, 35-37].

Overall, the occipital, temporal, and frontal lobes subserve the precise extract/categorization of the biological (i.e., living or not) and physical (i.e., bigger than a cat or not) features of the incoming spoken word under the modulation of the higher-order frontal lobe.

Other than the expected findings above, the hypothesized neural activity to be seen in the bilateral posterior superior temporal lobes failed to surpass the threshold of p(FWE) < .05 in the semantic > tone contrast. It is plausible that activity in these auditory regions was probably obscured by the tone baseline, as the posterior temporal speech areas are very likely to be intensely elicited by non-speech sounds such as the tones in the current study [28, 38, 39].

Unanticipated activations were also observed. The activation spotted in the left superior parietal lobule may occur due to visuospatial analysis that was needed in the semantic but not the tone condition, where the size of the object indicated by a heard noun had to be analysed and compared with the size of a cat. This implication is based on the role of this area in the visuospatial analysis of objects [40], visual search, and spatial selective attention [41].

It is also interesting to note semantic-irrelevant activations in the left sensorimotor cortex in the current study. The recruitment of this region is more likely to be attributed to the covert articulation of the words heard that were automatically elicited in the semantic but not the tone condition. Activity in the left sensorimotor cortex is unlikely due to motoric processes for button responses, in that the motoric activity evoked by the button-press appears to be cancelled out after contrasting the semantic condition with the tone baseline.

In addition to the cerebral activations stated above,

cerebellar activity in the right posterolateral cerebellum (Crus I/II) was also detected. Past empirical and meta-analytic studies have accumulated evidence for the preferential sensitivity of the anterior and posterior cerebellum to sensorimotor functions and higher-order cognitive processes, respectively [42-44]. Thus, the current observation of the posterolateral cerebellum in the semantics > tones contrast probably implies its involvement in the higher-order language functioning. Interestingly, in contrast to the left lateralization of the cerebral network underlying language processing, the present results showed cerebellar activity in the right hemisphere, reflecting crossed cerebro-cerebellar projections as reported by past studies [43-44]. In addition, the right Crus I/II has been suggested to subserve linguistic predictability, error adjustment, and phonological but not orthographic or semantic processing [43-45]. Thus, the current observation of the right Crus I/II activation was presumably associated with the phonological decoding of the word heard as well as participants' prediction of the characteristics of the upcoming word heard (e.g., whether it denotes a living thing bigger than a cat) but not directly correlated with semantic processing.

Overall, activations consistent with our hypothesis were found in the left occipito-temporal cortex, the left middle temporal gyrus, the right anterior superior temporal gyrus, the left ventral inferior frontal gyrus, and the left middle frontal gyrus, beyond the influence of task requirement. Due to sensitivity to task properties, minor inconsistencies are noted between our hypothesis and findings in the bilateral posterior superior temporal lobes, the left superior parietal lobule, the left sensorimotor cortex, and the right cerebellum (Crus I/II). These inconsistent regions are mostly situated in the dorsal stream except for the cerebellar regions [46-49], supporting the view that the degree of activity in the auditory dorsal stream is more sensitive to task variations compared to that in the auditory ventral stream [50].

Functional connectivity results

Functional connectivity was examined within a predefined network, in order to reveal potential interactions among the regions of the Chinese auditory lexicosemantic network, as well as to separate the auditory lexicosemantic processing of interest from irrelevant processing driven by task demands. Results showed that the semantic condition had greater connectivity than the tone condition among the right posterior superior temporal gyrus, the left middle frontal gyrus, and the left ventral inferior frontal gyrus. This indicates that an interconnected sub network among these three regions is related to the facilitation of the sensori-to-semantic transformation. It is believed that the right auditory cortex may be engaged in processing the complexity of acoustic features (e.g., pitch frequencies, spectral properties), which is greater in spoken words than tones. The left middle frontal gyrus is involved in integrating word form and meaning at the lexical level [28]. The neural activity observed in the right posterior superior temporal gyrus and the left middle frontal gyrus in tandem with the left ventral inferior frontal gyrus could be related to the retrieval of the target meaning from a spoken word with higher precision. These three regions may serve as

nodes to link and converge information from discrete brain sites, providing support for a Chinese auditory lexicosemantic network of highly-interconnected regions.

As reviewed earlier, the left posterior temporal lobe and the left frontal lobe have been demonstrated to be anatomically connected via the inferior fronto-occipital fasciculus to support the normal Chinese semantic functioning [8], and also functionally interconnected to subserve Chinese visual lexicosemantic processing [2, 6, 7]. In contrast to the intra-hemispheric connections reported in these previous studies, the current study observed connectivity between the left frontal lobe and the right posterior temporal lobe across the two hemispheres. The involvement of the right rather than the left posterior temporal lobe may still be linked to the non-linguistic acoustic differences between the current semantic and the tone conditions, such as the pitch frequencies (e.g., contour or level) and spectral properties (e.g., wide or narrow). According to the cortical asymmetry model [51], the right temporal lobe is more sensitive to certain aspects of the acoustic signal in comparison with the left temporal lobe. Therefore, the current observation of the inter-hemispheric connectivity between the left frontal lobe and the right posterior temporal lobe may be attributed to the fact that the task was presented aurally, and also due to considerable acoustic differences between the semantic and tone conditions. Nevertheless, the current finding serves as the primary evidence for the significance of the fronto-temporal connectivity in sensori-to-semantic transformation during Chinese lexicosemantic processing, specifically in the auditory domain.

The current results indicate that the left middle frontal gyrus and the left ventral inferior frontal gyrus could be core regions for Chinese auditory lexicosemantic processing. Although the right posterior superior temporal cortex was not significantly activated in the whole-brain functional map, it was functionally connected with the left frontal regions during auditory lexicosemantic processing. This is probably because acoustic processing alone may not be sufficient to facilitate the performance of the auditory semantic-tone task, where coordination of acoustic processing and higher-order semantic integration/selection is needed. Apparently, the inter-regional functional connectivity reveals some unique neural mechanisms, not observable through activation patterns. This implies the necessity to consider the function of a brain region within the neural network of which it is a part [28, 50], in order to further understand how different brain regions interact with one another in spoken Chinese word comprehension.

Significance of the study

Findings in the current study helped to deepen our understanding of how the linguistic nature of spoken Chinese words influences the neural mechanisms engaged. Based on the limited evidence available in the literature and our current investigations, a preliminary Chinese auditory lexicosemantic network is established for reference and to further improve future research.

In addition, the network based on the intact Chinese semantic system may serve as a baseline for us to further understand Chinese semantic disorders in dementia and develop potential interventions. For instance, if orthographic information does influence Chinese auditory semantic processing as revealed in the present study, it would be worth investigating whether the Chinese patients with alexia also have difficulty in listening comprehension in addition to the reading comprehension deficiency, or whether the neural basis of auditory semantic processing is different between the literate and the illiterate. Clinically, by integrating fMRI with other techniques such as Transcranial Magnetic Stimulation (TMS), we could further verify whether there is a need for neurosurgeons to consider the particularity of a Chinese neuro-network or the nuances between Chinese and English networks during the pre-surgical preparations for the Chinese-speaking population.

Limitations

Despite careful considerations of possible confounds, our study is not without limitations. Our task used a tone condition as the baseline. One could argue that the semantic > tone contrast may not purely target the auditory semantic network of interest given other differences between the semantic and tone conditions like pitch frequencies (e.g., contour or level) and spectral properties (e.g., wide or narrow). Meanwhile, the semantic and tone conditions also differed in whether the stimulus was linguistic (e.g., phonological processing involved or not). This semantic-tone task was adapted from Binder et al. [9], where the original task was devised to elicit auditory lexicosemantic network and it effectively controlled for the automatic non-linguistic processing such as involuntary sensorimotor activities, attentional control, and executive functioning in the tone baseline [9, 52-53]. Previous studies using this task paradigm also reliably activated the auditory lexicosemantic neuro-network [38, 39, 55, 56]. Nevertheless, future studies may consider designing a more comparable baseline in investigations of auditory lexicosemantic processing.

The missing behavioural data were less satisfactory as a result of technical difficulties, thereby preventing further analyses of behaviour-brain correlation and behaviour-brain regression, and ANCOVA with accuracy/latency as covariates could not be conducted. However, these analyses appear to be non-essential for the present study, where all participants were young adults with relatively low variability in performance. Furthermore, analysis of the remaining recorded trials showed that they were distributed throughout the entire run, partially confirming participants' maintenance of concentration and continued involvement during the task. There was also no significant difference of functional activation between the participants who had more missing behavioural data and those having less. Therefore, all the 23 participants were included for image analysis to ensure adequate sample size and statistical power.

Lastly, we only investigated functional connectivity in this study and were incapable of determining the directional influence from one region to another. Thus, the potential bottom-up and top-down processes could not be differentiated if based on the current connectivity results. We did not perform

effective connectivity analysis as it required a well-defined model for model comparisons. It is difficult to identify such a model considering the limited existing studies. Hopefully, more studies evaluating the Chinese auditory lexicosemantic network will be available in the future to help determine a well-defined network model to allow for effective connectivity exploration.

Conclusion

The current study verified the predicted activations in the left occipito-temporal cortex, the left middle temporal gyrus, the right anterior superior temporal gyrus, the left ventral inferior frontal gyrus, and the left middle frontal gyrus. The bilateral posterior superior temporal lobes that were hypothesized did not show significant activations. This is possibly obscured by the auditory processing of the tone baseline, as the bilateral auditory areas appear not to be linguistic-specific but are instead responsive to non-linguistic tones [28]. Nevertheless, co-activity between the right posterior superior temporal cortex and frontal areas was revealed in functional connectivity results, indicating the involvement of the posterior superior temporal cortex in Chinese auditory lexicosemantic processing.

In summary, the semantic network of spoken Chinese words is thought to involve the bilateral posterior superior temporal lobes, the left middle frontal gyrus, the left ventral inferior frontal gyrus, the anterior superior temporal cortex, the left middle temporal gyrus, and the left occipito-temporal cortex. Within this Chinese auditory lexicosemantic network, the right posterior superior temporal gyrus, the left middle frontal gyrus, and the left ventral inferior frontal gyrus were identified as an interconnected sub network through functional connectivity analysis. This verifies the significance of the fronto-temporal connectivity in the Chinese lexicosemantic network specifically in the auditory domain.

The current Chinese language network appears to be generally consistent with alphabetic writing systems, with some minor specificities noted. The auditory lexicosemantic processing of Chinese appears less relevant to the left inferior parietal lobule underlying the sub-lexical-level processing while more associated with the lexical-level processing in the left middle frontal gyrus. In addition, the occipito-temporal visual cortex seems to be involved in spoken Chinese lexicosemantic processing, indicating the involvement of orthography in Chinese auditory lexicosemantic processing despite the absence of visually-presented characters. From the linguistic perspective, these findings likely reflect that Chinese characters, being evolved from semantic-represented drawings, are inherently more advantageous over corresponding spoken forms in terms of supplying semantic clues and differentiating homophones during the lexical access of Chinese [57].

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

HL contributed to data analysis and manuscript writing. AC obtained funding, designed the study, oversaw the data collection, and contributed to the overall manuscript write-up. Both authors reviewed and approved the manuscript.

Acknowledgments

References

1. Liu L, Deng X, Peng D, Cao F, Ding G, et al. 2009. Modality- and task-specific brain regions involved in Chinese lexical processing. *J Cogn Neurosci* 21(8): 1473-1487. https://doi.org/10.1162/jocn.2009.21141

2. Wu X, Lu J, Chen K, Long Z, Wang X, et al. 2009. Multiple neural networks supporting a semantic task: an fMRI study using independent component analysis. *Neuroimage* 45(4): 1347-1358. https://doi.org/10.1016/j.neuroimage.2008.12.050

3. Xiao Z, Zhang JX, Wang X, Wu R, Hu X, et al. 2005. Differential activity in left inferior frontal gyrus for pseudowords and real words: an event-related fMRI study on auditory lexical decision. *Hum Brain Mapp* 25(2): 212-221. https://doi.org/10.1002/hbm.20105

4. Zou L, Packard JL, Xia Z, Liu Y, Shu H. 2016. Neural correlates of morphological processing: Evidence from Chinese. *Front Hum Neurosci* 9: 714. https://doi.org/10.3389/fnhum.2015.00714

5. James CT. 1975. The role of semantic information in lexical decisions. *J Exp Psychol Hum Percept Perform* 1(2): 130-136. https://psycnet.apa.org/doi/10.1037/0096-1523.1.2.130

6. Fan LY, Chou TL. 2012. Hierarchical model comparisons on effective connectivity in semantic judgments of Chinese characters. *NTU Scholars* 54(1): 31-46.

7. Fan LY, Lee SH, Chou TL. 2010. Interaction between brain regions during semantic processing in Chinese adults. *Language and Linguistics* 11(1): 159-182.

8. Han Z, Ma Y, Gong G, He Y, Caramazza A, et al. 2013. White matter structural connectivity underlying semantic processing: evidence from brain damaged patients. *Brain* 136(Pt 10): 2952-2965. https://doi.org/10.1093/brain/awt205

9. Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, et al. 1997. Human brain language areas identified by functional magnetic resonance imaging. *J Neurosci* 17(1): 353-362.

10. Ashburner J. 2007. A fast diffeomorphic image registration algorithm. *Neuroimage* 38(1): 95-113. https://doi.org/10.1016/j.neuroimage.2007.07.007

11. Sladky R, Friston KJ, Tröstl J, Cunnington R, Moser E, et al. 2011. Slice-timing effects and their correction in functional MRI. *Neuroimage* 58(2): 588-594. https://doi.org/10.1016/j.neuroimage.2011.06.078

12. Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, et al. 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 46(3): 786-802. https://doi.org/10.1016/j.neuroimage.2008.12.037

13. Yassa MA, Stark CE. 2009. A quantitative evaluation of cross-participant registration techniques for MRI studies of the medial temporal lobe. *Neuroimage* 44(2): 319-327. https://doi.org/10.1016/j.neuroimage.2008.09.016

14. Liu H. 2017. Investigating the behavioral and neural correlates of auditory semantic processing in the Chinese language. Doctoral Thesis, Nanyang Technological University, Singapore.

15. Brett M, Anton JL, Valabregue R, Poline JB. 2002. Region of interest analysis using an SPM toolbox. 8th International Conference on Functional Mapping of the Human Brain, Neuroimage, Japan.

16. Friston K. 2002. Beyond phrenology: what can neuroimaging tell us about distributed circuitry? *Annu Rev Neurosci* 25(1): 221-250. https://doi.org/10.1146/annurev.neuro.25.112701.142846

17. Stevens MC. 2009. The developmental cognitive neuroscience of functional connectivity. *Brain Cognition* 70(1): 1-12. https://doi.org/10.1016/j.bandc.2008.12.009

18. Fingelkurts AA, Fingelkurts AA, Kähkönen S. 2005. Functional connectivity in the brain--is it an elusive concept? *Neurosci Biobehav Rev* 28(8): 827-836. https://doi.org/10.1016/j.neubiorev.2004.10.009

19. Sala-Llonch R, Bartrés-Faz D, Junqué C. 2015. Reorganization of brain networks in aging: a review of functional connectivity studies. *Front Psychol* 6: 663. https://doi.org/10.3389/fpsyg.2015.00663

20. Smith SM, Vidaurre D, Beckmann CF, Glasser MF, Jenkinson M, et al. 2013. Functional connectomics from resting-state fMRI. *Trends Cogn Sci* 17(12): 666-682. https://doi.org/10.1016/j.tics.2013.09.016

21. Whitfield-Gabrieli S, Nieto-Castanon A. 2012. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect* 2(3): 125-141. https://doi.org/10.1089/brain.2012.0073

22. Eklund A, Nichols T, Knutsson H. 2015. Can parametric statistical methods be trusted for fMRI based group studies? *PNAS* 113(28): 7900-7905.

23. Eklund A, Nichols TE, Knutsson H. 2016. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci U S A* 113(28): 7900-7905. https://doi.org/10.1073/pnas.1602413113

24. Brodmann's Interactive Atlas. [http://www.fmriconsulting.com/brodmann/Introduction.html]

25. Cohen L, Lehéricy S, Chochon F, Lemer C, Rivaud S, et al. 2002. Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain* 125(Pt 5): 1054-1069. https://doi.org/10.1093/brain/awf094

26. Harley TA. 2014. The Psychology of Language: From Data to Theory. Psychology Press, Washington.

27. Kosslyn SM, Pascual-Leone A, Felician O, Camposano S, Keenan JP, et al. The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science* 284(5411): 167-170. https://doi.org/10.1126/science.284.5411.167

28. Price CJ. 2012. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage* 62(2): 816-847. https://doi.org/10.1016/j.neuroimage.2012.04.062

29. Wise RJ, Price CJ. 2006. Functional neuroimaging of language. In: Cabeza R, Kingstone A (Eds), Handbook of Functional Neuroimaging of Cognition, pp 191-228.

30. Wu CY, Ho MH, Chen SH. 2012. A meta-analysis of fMRI studies on Chinese orthographic, phonological, and semantic processing. *Neuroimage* 63(1): 381-391. https://doi.org/10.1016/j.neuroimage.2012.06.047

31. Binder JR, Desai RH, Graves WW, Conant LL. 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex* 19(12): 2767-2796. https://doi.org/10.1093/cercor/bhp055

32. Friederici AD. 2012. The cortical language circuit: from auditory perception to sentence comprehension. *Trends Cogn Sci* 16(5): 262-268. https://doi.org/10.1016/j.tics.2012.04.001

33. Herrmann B, Maess B, Hahne A, Schröger E, Friederici AD. 2011. Syntactic and auditory spatial processing in the human temporal cortex: an MEG study. *Neuroimage* 57(2): 624-633. https://doi.org/10.1016/j.neuroimage.2011.04.034

34. Farahibozorg S, Woollams AM, Cooper E, Evans G, Chen Y, et al. 2017. Semantic Hub or Convergence Zones? EEG/MEG Evidence for a Central Role of ATL in Semantic Processing. Organization for Human Brain Mapping Annual Conference.

35. Lee SH, Booth JR, Chen SY, Chou TL. 2011. Developmental changes in the inferior frontal cortex for selecting semantic representations. *Dev Cogn Neurosci* 1(3): 338-350. https://doi.org/10.1016/j.dcn.2011.01.005

36. Liuzzi AG, Bruffaerts R, Peeters R, Adamczuk K, Keuleers E, et al. 2017. Cross-modal representation of spoken and written word meaning in left pars triangularis. *Neuroimage* 150: 292-307. https://doi.org/10.1016/j.neuroimage.2017.02.032

37. Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, et al. 1999. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage* 10(1): 15-35. https://doi.org/10.1006/nimg.1999.0441

38. Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Rao SM, et al. 1999. Conceptual processing during the conscious resting state. A functional MRI study. *J Cogn Neurosci* 11(1): 80-95.

39. Cabeza R. Kingstone A. 2006. Handbook of Functional Neuroimaging of Cognition. MIT Press, London.

40. Haxby JV, Grady CL, Horwitz B, Ungerleider LG, Mishkin M, et al. 1991. Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proc Natl Acad Sci U S A* 88(5): 1621-1625. https://doi.org/10.1073/pnas.88.5.1621

41. Booth JR, Burman DD, Meyer JR, Lei Z, Trommer BL, et al. 2003. Neural development of selective attention and response inhibition. *Neuroimage* 20(2): 737-751. https://doi.org/10.1016/S1053-8119(03)00404-X

42. Chen SH, Desmond JE. 2005. Temporal dynamics of cerebro-cerebellar network recruitment during a cognitive task. *Neuropsychologia* 43(9): 1227-1237.

43. E KH, Chen SH, Ho MH, Desmond JE. 2014. A meta-analysis of cerebellar contributions to higher cognition from PET and fMRI studies. *Hum Brain Mapp* 35(2): 593-615. https://doi.org/10.1002/hbm.22194

44. Stoodley CJ, Schmahmann JD. 2009. Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. *Neuroimage* 44(2): 489-501. https://doi.org/10.1016/j.neuroimage.2008.08.039

45. Lesage E, Hansen PC, Miall RC. 2017. Right lateral cerebellum represents linguistic predictability. *J Neurosci* 37(26): 6231-6241. https://doi.org/10.1523/JNEUROSCI.3203-16.2017

46. Hickok G, Poeppel D. 2000. Towards a functional neuroanatomy of speech perception. *Trends Cogn Sci* 4(4): 131-138.

47. Rauschecker JP. 1998. Cortical processing of complex sounds. *Curr Opin Neurobiol* 8(4): 516-521. https://doi.org/10.1016/s0959-4388(98)80040-8

48. Scott SK, Wise RJ. 2004. The functional neuroanatomy of prelexical processing in speech perception. *Cognition* 92(1-2): 13-45. https://doi.org/10.1016/j.cognition.2002.12.002

49. Wise RJ, Scott SK, Blank SC, Mummery CJ, Murphy K, et al. 2001. Separate neural subsystems within Wernicke's area. *Brain* 124(1): 83-95. https://doi.org/10.1093/brain/124.1.83

50. Hickok G, Poeppel D. 2007. The cortical organization of speech processing. *Nat Rev Neurosci* 8(5): 393-402. https://doi.org/10.1038/nrn2113

51. McGettigan C, Scott SK. 2012. Cortical asymmetries in speech perception: what's wrong, what's right and what's left? *Trends Cogn Sci* 16(5): 269-276. https://doi.org/10.1016/j.tics.2012.04.006

52. Binder JR, Rao SM, Hammeke TA, Frost JA, Bandettini PA, et al. 1995. Lateralized human brain language systems demonstrated by task subtraction functional magnetic resonance imaging. *Arch Neurol* 52(6): 593-601. https://doi.org/10.1001/archneur.1995.00540300067015

53. Démonet JF, Chollet F, Ramsay S, Cardebat D, Nespoulous JL, et al. 1992. The anatomy of phonological and semantic processing in normal subjects. *Brain* 115(Pt 6): 1753-1768. https://doi.org/10.1093/brain/115.6.1753

54. Yang FG, Edens J, Simpson C, Krawczyk DC. 2009. Differences in task demands influence the hemispheric lateralization and neural correlates of metaphor. *Brain Lang* 111(2): 114-124. https://doi.org/10.1016/j.bandl.2009.08.006

55. Perani D, Dehaene S, Grassi F, Cohen L, Cappa SF, et al. 1996. Brain processing of native and foreign languages. *Neuroreport* 7(15-17): 2439-2444. https://doi.org/10.1097/00001756-199611040-00007

56. Price CJ, Wise RJ, Frackowiak RS. 1996. Demonstrating the implicit processing of visually presented words and pseudowords. *Cereb Cortex* 6(1): 62-70.

57. Perfetti CA, Dunlap S. 2008. Universal Principles And Writing System Variations In Learning To Read. Learning To Read Across Languages. Mahwah, NJ: Erlbaum.