

ORIGINAL RESEARCH

The Neural Correlates of Spoken Sentence Comprehension in the Chinese Language: An fMRI Study

This article was published in the following Dove Press journal: Psychology Research and Behavior Management

Hengshuang Liu 1.2
SH Annabel Chen 2-5

¹Bilingual Cognition and Development Lab, National Key Research Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou, People's Republic of China; ²Psychology, School of Social Sciences (SSS), Nanyang Technological University, Singapore; ³Centre for Research and Development in Learning (CRADLE), Nanyang Technological University, Singapore; ⁴Lee Kong Chian School of Medicine (LKCMedicine), Nanyang Technological University, Singapore; ⁵National Institute of Education, Nanyang Technological University, Singapore

Purpose: Everyday social communication emphasizes speech comprehension. To date, most neurobiological models regarding auditory semantic processing are based on alphabetic languages, where the character-based languages such as Chinese are largely underrepresented. Thus, the current study attempted to investigate the neural network of speech comprehension specifically for the Chinese language.

Methods: Twenty-two native Mandarin Chinese speakers were imaged while performing a passive listening task of forward and backward sentences. Sentences were used as task stimuli, as sentences compared with words were more frequently utilized in daily speech comprehension. **Results:** Our results suggested that spoken Chinese sentence comprehension may involve a neural network comprising the left middle temporal gyrus, the left anterior temporal lobe, and the bilateral posterior superior temporal lobes. The occipitotemporal visual cortex was not found to be significantly involved with the sentence-level network of spoken Chinese comprehension, as bottom-up visualization process from homophones to visual forms may be less needed due to the availability of top—down contextual controls in sentence processing. In addition, no significant functional connectivity was observed, likely obscured by the low cognitive demand of the task conditions. Limitations and future directions were discussed.

Conclusion: The current Chinese network seems to largely resemble the auditory semantic network for alphabetic languages but with features specific to Chinese. While the left inferior parietal lobule in the dorsal stream may have little involvement in the listening comprehension of Chinese sentences, the ventral neural stream via the temporal cortex appears to be more emphasized. The current findings deepen our understanding of how the semantic nature of spoken Chinese sentences influences the neural mechanism engaged.

Keywords: Chinese, character-based languages, auditory semantic network, spoken sentences, speech comprehension

Introduction

Everyday interpersonal communication mostly emphasizes speech comprehension. Speech is usually relayed in sentences, which comprise sets of words but convey meanings far more abundant than the simple sum of the constituent words. Spoken sentence comprehension thus relies not only on the identification of individual spoken words, but also on the contextual information giving rise to the expected meanings. These meanings of individual spoken words in a sentence are determined by the words preceding and following it, and the frequency of such occurrence. In a Chinese sentence, the lack of grammatical inflections such as word category, case, number,

Correspondence: SH Annabel Chen Email annabelchen@ntu.edu.sg

and person therefore places greater reliance on the word order in Chinese sentence comprehension.^{2,3} Every word in a sentence must be placed in an appropriate sequence to be compatible with the remaining words in the sentence; changing one word or the position of the word sometimes may completely alter the sentence meaning. 4-6 It is thus interesting to understand how the brain allows a Chinese listener to sequentially bind the isolated words into a coherent sentence so that additional information not contained in the single words could be delivered.

To date, most neurobiological models of auditory semantic processing are based on alphabetic languages like English, where character languages like Chinese are largely underrepresented. One of the most prevailing models is the dorsaland ventral-stream model. ^{7–11} This dual-stream model entails a specific ventral stream interfacing sounds/prints with semantic representations, which involves the left anterior ventral Broca's area (BA45/47), the left anterior superior temporal gyrus, the left middle temporal gyrus, and the left occipitotemporal cortex. In addition, the model also conceptualizes the ventral stream as less left-lateralized and more bilaterally organized relative to the strongly left-dominant dorsal stream.

In Price's anatomical model, 12 a brain map of auditory semantic areas with their most consistent functions has been depicted. Based on this model, auditory semantic processing usually activates left-lateralized areas. The anterior and posterior portions of the superior temporal lobe are specialized for semantic composition and phonological perception, respectively. Activations for sentences normally spread anteriorly to the temporal pole. Posterior middle/inferior temporal activations are regulated by task demands. Ventral inferior frontal areas (pars orbicularis and pars triangularis) may subserve the selection of task-related semantic attributes. The lateral parietal areas such as the angular gyri are involved in the cross-modal integration of semantic features.

In contrast to these models, the cortical asymmetry model specifically focuses on speech perception rather than speech comprehension.¹³ According to this model, the left and right temporal lobes are preferentially sensitive to linguistic-specific cues and non-linguistic acoustic signals, respectively.

Based on the semantic-related regions identified in these models, a possible auditory semantic neural network for alphabetic languages is likely to include the left inferior parietal lobule, the left anterior superior temporal gyrus, the bilateral posterior superior temporal lobes, the

left ventral inferior frontal gyrus, the left middle temporal gyrus, and the left occipitotemporal cortex.

It is unclear whether and how this network applies to the auditory semantic processing of the character-based Chinese language, given the linguistic differences between Chinese and English: While the semantic representation in English seems to be more easily retrieved from the spoken than written form, visual Chinese forms are presumably less ambiguous and more available to semantic access compared to auditory forms, due to the existence of pervasive homophones and semantic radicals in the Chinese language.

Based on past studies of visual Chinese recognition, reading Chinese involved weaker activation in the left inferior parietal lobule and greater activation in the left middle frontal gyrus as compared with English reading. This is likely due to the greater orthographic arbitrariness in Chinese than English. 14-20 While the left inferior parietal lobule is usually related to sound assembly from the constituent phonemes, the function of assigning a syllable to an ideographic character is normally mediated by the left middle frontal gyrus.

However, little is known about the spoken Chinese neuro-network, as a literature search found that only five neuroimaging studies evaluated spoken Chinese comprehension. Four of these five studies investigated Chinese speech comprehension using word stimuli. 21-24 The earliest study among these four found that the bilateral occipitaltemporal cortices (BA37) and the bilateral middle temporal gyri (BA21) had greater activity when making lexical decision to spoken disyllabic Chinese words (eg, '太阳'/tai4 yang2/sun) contrasted to making lexical decision to spoken pseudowords (eg, '领村' /ling3 cun1/ nonword without meaning).²³ 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, as was the case in the spoken recognition of alphabetic languages. 25,26 In addition to the bilateral visual cortices, the bilateral middle temporal gyri were also seen in the words > pseudowords contrast. This possibly indicates the greater lexicosemantic representations in real words than pseudowords.

More recently, similar findings are observed in two studies using similar task paradigms, where auditory meaning relatedness judgment compared to tone baseline

evoked greater activations in the bilateral posterior temporal lobes (BA22 in Liu et al, 2009; BA48 in Zou et al, 2015), the left occipitotemporal cortex (BA18 in Liu et al, 2009; BA37 in Zou et al, 2015), and the left ventral frontal lobe (BA47 in Liu et al, 2009; BA45 in Zou et al, 2015). Ba45 in Zou et al, 2015). Chinese auditory lexicosemantic processing. It has also been interpreted in these two studies that the recruitment of the left occipitotemporal cortex was driven by the interaction of orthographic and phonological representations during spoken Chinese word recognition. However, the regions of interest (ROIs) applied in Zou et al's study (2015) were defined from the same study, thus limiting power of inference of the findings.

In addition to regional activation pattern, interregional connectivity network underpinning Chinese auditory lexicosemantic processing was examined in Wu et al's study (2009) using multivariate independent component analysis (ICA).²² While the regional activation results seen in the occipitotemporal cortex (right lingual gyrus, BA18) and the left ventral frontal lobe (BA45/47) were consistent with Liu et al's (2009) and Zou et al's (2015), several intensely-connected networks were identified within the extensive fronto-temporal cortex when contrasting Chinese auditory semantic dangerousness judgment (eg, '手枪'/shou3 giang1/gun) to a rest 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 (2009) 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.

In addition to these four studies examining spoken Chinese comprehension using word stimuli, only one published study was found to assess Chinese auditory semantic neural networks using sentence stimuli.²⁷ In fact, sentences are more frequently utilized in everyday speech comprehension and thus worthy of further research investigations.

In Xu et al's study (2013) listening to the scrambled sentence composed by randomly-selected words compared

with the sentence consisting of misplaced-consonant syllables elicited stronger activations in the left occipitotemporal cortex (BA37/20), the left middle temporal gyrus (BA21), and the left ventral inferior frontal gyrus (BA47), which is partly consistent with the auditory semantic neuro-network for alphabetic languages. However, this activation pattern is essentially more representative for lexicosemantic processing rather than for sentence-level semantic processing, despite the utilization of sentence-level stimuli in this study. The aim of Xu et al's study (2013) was to investigate whether lexical meaning can be accessed in pitch-flattened (monotone) sentences in tonal languages like Mandarin Chinese. Therefore, the contrast of scrambled sentences and misplaced consonant sentences was used as a 'localizer' to target regions responsible for accessing meaning to words but not to sentences, assuming word intelligibility was the major difference between these two types of sentences: Scrambled sentences were syntactically anomalous and unintelligible at the sentence level, while the consonant replacement led to syntactic anomaly and unintelligibility at the levels of both words and sentences. Given the actual differences between the two conditions, the neural networks underlying spoken Chinese sentence comprehension still remains unanswered.

There is limited literature about language networks for spoken Chinese sentence comprehension, and even less knowledge is available for the interregional connectivity pattern underpinning this process. To date, anatomical connectivity underlying the Chinese semantic system has been examined using diffusion tensor imaging.²⁸ In this study, reduced fractional anisotropy and greater lesion percentage in the left inferior fronto-occipital fasciculus were found to be 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 grey matter involvement. It is not far-fetched to presume that the underlying anatomical networks for Chinese semantic processing also support Chinese auditory sentence comprehension. If so, the semantic processing of spoken Chinese sentences would likely involve the left inferior fronto-occipital fasciculus connecting the posterior lateral temporal cortex to frontal cortex as indicated for Chinese semantic comprehension.

As with anatomical connectivity, only one study was found to investigate functional connectivity underlying Chinese sentential semantic processing.²⁹ However, sentence stimuli in this study were presented visually instead

of aurally. Results showed that the incongruent > congruent contrast produced stronger functional connectivity between the left/right ventral inferior frontal gyrus and several inhibitory-related regions such as BA40 (inferior parietal lobule) and BA46 (dorsal lateral prefrontal cortex), suggesting that the comprehension system attempted to suppress or replace the semantically-violated word in a given incongruent sentence (eg,'连这么大的声音张都能 听清楚, 太敏锐了。'/lian2 zhe4 me da4 de sheng1 yin1 zhang1 dou1 neng2 ting1 qing1 chu3, tai4 min3 rui4 le/ Even such a loud sound can be heard by Zhang; he has a sharp hearing). However, the authors did not examine congruent > incongruent conditions, which could provide more information regarding semantic comprehension without conflict. Nevertheless, the finding from this study is still valuable for the understanding of the functional connectivity specific to Chinese sentence comprehension.

Given the limited investigation into the semantic neural networks underlying spoken Chinese sentence processing, the current study was conducted to fill this research gap. We used an archival fMRI data set collected from native Mandarin Chinese participants performing a forwardbackward passive listening task. This task paradigm allows us to identify the sentential semantic component through the forward > backward contrast, since the sentence-level semantic information that was intact in the forward condition was largely eliminated in the backward baseline. Based on the auditory semantic network for alphabetic languages and past studies of Chinese neuro-networks, whole-brain activations were hypothesized to be observed in 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 occipitotemporal cortex. To further identify the functional connections between these hypothesized regions, a priori ROIs were created, and functional connectivity between these ROIs was computed.

Materials and Methods

Participants

An archival neuroimaging data set of 26 native speakers of Mandarin Chinese was employed in the current study. The study was conducted in Taiwan and all participants used traditional Chinese characters in their everyday 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, and this study was conducted in accordance with the Declaration of Helsinki.

Three male participants were excluded due to the poor image quality. Another male participant was excluded due to missing data from the follow-up sentence comprehension test and his engagement in the passive listening task could not be verified. The remaining 22 right-handed participants were included in subsequent data analysis (12 females; mean age = 25.7 years, SD = 4.3 years, range = 19–35 years; mean handedness score = 87.2, SD = 17.8, range = 42.8–100).

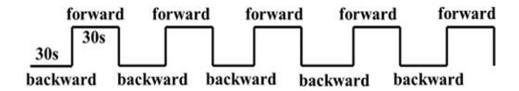
Procedure

A block design passive listening task was adapted from Maldjian et al.³⁰ In the task, participants were instructed to listen to a short story played for 30 sec in the forward block, and listened to the same text played backwards for another 30 sec in the backward block. It is of note that the audio in the backward condition was a complete reversal of the forward audio. Thus, the backward audio was not comprehensible at either the lexical or the sentential level. Five forward blocks and five backward blocks were alternated in one run (Figure 1), and the task was comprised of five runs in total. The task was considered to provide a well-matched between-condition comparison, as both forward and backward conditions had equivalent spectral information profile such as intensity, pitch, and amplitude (Table 1) but mainly differed in their meaningfulness. Although no response was required for the task, participants' concentration on the task was assessed by a followup comprehension test after the scan, which included 10 questions requiring responses of "right" or "wrong" with no response time-limit.

Image Acquisition

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

A The condition sequence in a task run



B Examples of a forward and a backward block

A forward block

Chinese 农场里住着一只黑猫,小主人很喜欢它。黑猫渐渐长大,它总嫌自己的毛太黑,鼻子太塌,尾巴太长,叫声又难听,总因为羡慕而想模仿别人,所以大家叫它"模仿猫"。

Pinyin /nong2 chang3 li3 zhu4 zhe yi4 zhi1 hei1 mao1, xiao3 zhu3 ren2 hen3 xi3 huan1 ta1. Hei1 mao1 jian4 jian4 zhang3 da4, ta1 zong3 xian2 zi4 ji3 de mao2 tai4 hei1, bi2 zi tai4 ta1, yi3 ba1 tai4 chang2, jiao4 sheng1 you4 nan2 ting1, zong3 yin1 wei2 xian4 mu4 er2 xiang3 mo2 fang3 bie2 ren2, suo3 yi3 da4 jia1 jiao4 ta1 "mo2 fang3 mao1"/

English There is a black cat living in the farm, and his little master likes him very much. As the cat grows up, he always feels inferior because of his black fur, flat nose, long tail, and unpleasant meow. He is thus envious of others and try to imitate others, and therefore he is called "mimicker".

A backward block: a reversal of the forward audio, thus completely unintelligible

Figure I The design of the forward-backward passive listening task.

Participants were imaged 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 preprocessing protocol implemented in Statistical Parametric Mapping 8 (SPM8; Department of Cognitive Neurology, London, UK).³¹ The first stage was conducted for each individual 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,³² a realignment of the functional images to the first volume, a co-registration of the T1 MPRAGE image

Table I The Acoustic Characteristics of the Forwardly- and Backwardly-Presented Audio Stimuli

		Forward (n=25 Blocks)	Backward (n=25 Blocks)	Paired-Sample Difference
Intensity (dB)	Min	25.6±5.2	25.5±5.2	t (24) = 1.265, p = 0.218
	Max	73.7±2.4	73.7±2.4	t (24) = 0.912, p = 0.371
	Mean	62.7±1.9	62.7±1.9	t (24) = -0.979, p = 0.337
Pitch (Hz)	Min	76.0±6.4	76.3±6.5	t (24) = -1.367, p = 0.184
	Max	447.8±66.4	455.3±63.1	t (24) = -1.180, p = 0.249
	Mean	209.1±5.2	209.2±5.2	t (24) = -1.084, p = 0.289
Amplitude (Pascal)	Min	-0.259±0.067	-0.259±0.067	t (24) = 0, p = 1.000
	Max	0.268±0.069	0.268±0.069	t (24) = 0, p = 1.000
	Mean	0.0233±0.0002	0.0233±0.0002	t (24) = -0.388, p = 0.702

Note: The above phonetic information was extracted using the Praat toolbox (https://www.fon.hum.uva.nl/praat/).

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 image 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 preprocessing stage went through conventional preprocessing, the second and third stages underwent Diffeomorphic Anatomical RegisTration using Exponentiated Lie algebra (DARTEL) preprocessing instead for improved normalization. 33,34

After the three stages of preprocessing, whole-brain activation maps in the forward > backward contrast were obtained for each participant. Group activations across all participants were computed in the forward > backward contrast using a one-sample t-test. Results were thresholded at p-FWE <0.05, cluster size \geq 10 voxels. The FWE correction has been validated to generate more robust results as compared to the FDR-corrected or uncorrected results. By combining a voxel-level threshold (p-FWE < 0.05) with the cluster-based control (\geq 10 voxels), the overly small clusters likely resulting from false positives could be excluded, allowing for the suprathreshold clusters to be specifically identified with stronger statistical power. 35,36

In addition to the whole-brain analyses, ROIs were also created for the examination of possible functional connections between them. Based on our hypothesis, 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 occipitotemporal cortex were selected as a priori ROIs. The center coordinates of these ROIs were identified from a meta-analysis conducted by the first author (see Table 2 for more information).³⁷ Using the MarsBaR toolbox, 38 the ROIs of the left posterior superior temporal gyrus, the right posterior superior temporal gyrus, the left anterior superior temporal gyrus, the left occipitotemporal cortex, the left middle frontal gyrus, the left ventral inferior frontal gyrus, and the left middle temporal gyrus were defined as 8mm spheres centered at (-58, -14, -2), (64, -6, -6), (-50, 14, -16), (-56, -54,-14), (-48, 24, 24), (-42, 36, -4), and (-60, -42, -2), respectively.

Functional connectivity between every pair of these predefined ROIs was calculated in the forward > backward contrast by CONN 17b (https://www.nitrc.org/pro jects/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. 39,40 When distributed brain ROIs display strongly correlated patterns of neural activity change, it is taken as evidence that these ROIs are functionally connected. 41 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

Table 2 Liu's Meta-Analytic Results (2017) in (a) Chinese & English Visual Semantic Processing and (b) English Auditory Semantic Processing

Volume	Region	ВА	MNI Coordinate			Max ALE
(mm³)			х	у	z	
(a) Chinese &	& English visual semantic processing (n = 125 studies)	•	•	•	
5848	Left middle frontal gyrus Left inferior frontal gyrus	46 9, 44–47	-48 -42	24 36	24 -4	0.028 0.02
2024	Left superior frontal gyrus Left medial frontal gyrus	8	-2 0	24 8	48 52	0.038 0.015
1136	Left middle temporal gyrus	21	-60	-42	-2	0.02
(b) English a	uditory semantic processing (n = 40 s	tudies)	'	'	•	<u> </u>
8320	Left superior temporal gyrus	21, 41	-58	-14	-2	0.044
6320	Left inferior frontal gyrus Left middle frontal gyrus Left insula	47 46 13	-46 -50 -36	28 24 22	-6 24 14	0.033 0.021 0.021
4576	Right superior temporal gyrus Right middle temporal gyrus	22 21	64 64	-6 -30	-6 0	0.038 0.021
3016	Left inferior temporal gyrus Left fusiform gyrus	20 37	-56 -46	-54 -58	-14 -12	0.023 0.019
1968	Left superior temporal gyrus	38	-50	14	-16	0.026
1336	Left supramarginal gyrus Left middle temporal gyrus	40 39	-54 -42	-52 -60	28 24	0.018 0.016

Notes: The highlighted coordinates were defined as the ROI centers in the current study; ALE, Activation-Likelihood Estimation; Meta-analysis could not be directly conducted for the Chinese auditory semantic processing of interest due to the overly small number of Chinese auditory semantic studies (n = 5) compared to the required minimal dataset size of 17 studies in a meta-analysis (Eickhoff et al, 2016); Individual ALE maps were thresholded at cluster level p < 0.05, with clusters generated under 1000 permutations and a threshold of uncorrected p < 0.001; Contrast analyses were thresholded under a threshold of p < 0.05 (uncorrected) with 10,000 permutations and a minimum cluster volume of 250 mm³.

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. 42,43 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 interscan reliability.⁴⁴ Results of functional connectivity were thresholded at p-uncorrected (connection-level) <0.001 and network based statistics (NBS) p-FWE (by intensity) <0.05. NBS is a non-parametric approach based on permutation tests that looks at the extent of specific subnetworks of interconnected ROIs. It was verified that stronger power could be gained by combining a connection-level threshold with NBS thresholding relative to the individual connection-level inference.⁴⁵

Results

In the comprehension test that followed the passive listening task, the 22 participants had a mean accuracy of 86% (SD = 8%, range = 70–100%) and a mean reaction time of 2863 ms (SD = 2593 ms, range = 1413–12,504 ms). This accuracy ascertained participants' attentiveness and engagement in the passive listening task, so that the neuroimaging results are supposed to reflect the neural network that supports speech comprehension rather than simple sound perception.

Table 3 and Figure 2 show the whole-brain activity maps of the 22 participants, where greater activations for the forward than the backward condition were observed in the bilateral superior temporal gyri (BA38), the right superior temporal gyrus (BA22), the left middle temporal gyrus (BA21), the left supplementary motor area (BA6), and the left precentral gyrus (BA6). Corresponding brain regions and Brodmann areas (BAs) of the resulted coordinates were identified using the respective templates in the

Table 3 Peak Coordinates Within the Significant Clusters of the Whole-Brain Activation in the Forward > Backward Contrast

Region	ВА	MNI			Cluster size	Т
		x	у	z		
Left superior temporal gyrus	38	-57	8	-15	283	11.53
Left middle temporal gyrus	21	-65	-8	-11	20	9.65
Right superior temporal gyrus	38	57	11	-19	31	8.18
Right superior temporal gyrus	22	57	-19	-4	24	7.72
Left precentral gyrus	6	-42	0	46	16	7.97
Left supplementary motor area	6	-8	8	61	13	8.67

Notes: Cluster criterion, p-FWE < 0.05, cluster size ≥ 10 voxels.

Abbreviation: BA, Brodmann area.

MRIcron toolbox (https://www.nitrc.org/projects/mricron), the MARINA toolbox, 46 and the Yale BioImage Suite Online (https://bioimagesuiteweb.github.io/webapp/mni2 tal.html).

No significant functional connectivity was observed between any pair of ROIs in the forward > backward contrast.

Discussion

The current study aimed to identify the neural networks of spoken Chinese sentence comprehension using a forward-backward passive listening task. Based on the past literature, the Chinese auditory sentential semantic network was hypothesized to involve 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 whole-brain results validated that most of the expected regions were involved except for the left frontal regions and the left occipitotemporal cortex.

Whole-Brain Activation Results

In the forward > backward contrast, expected activations were observed in the left middle temporal gyrus (BA21), the left anterior superior temporal gyrus (BA38), and the

right posterior superior temporal gyrus (BA22). No significant activity was found in the hypothesized regions of the left middle frontal gyrus, the left ventral inferior frontal gyrus, and the left occipitotemporal cortex, while unexpected activations were noted in the left sensorimotor area (BA6).

The left middle temporal gyrus observed in the forward > backward contrast probably reflected the greater lexicosemantic processing involved in the forward than the backward condition, given the role of this region in lexicosemantic representations.⁴⁷

In contrast to the middle temporal gyrus which underpins lexicosemantic storage, the left anterior temporal lobe (BA38) is usually involved in grouping words into a larger unit for sentence-level comprehension. 48,49 In a meta-analysis of 164 semantic-related neuroimaging studies, 50 the anterior temporal lobe has been verified to be steadfastly activated in combinational semantic representations, above and beyond the influences of stimuli or tasks (eg, passive listening, passive reading, semantic judgment). In addition to the function of sentential semantic composition, the current involvement of the left anterior temporal lobe (BA38) may also be related to its role in building basic syntactic structure such as quantifier or prepositional phrases. 10,51–55 Anatomically, a segregation has been reported in the left anterior temporal lobe, with its

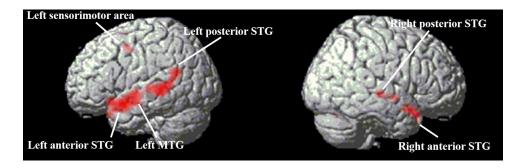


Figure 2 The whole-brain activation maps in the forward > backward contrast; p-FWE < 0.05, cluster size ≥ 10 voxels. Abbreviations: MTG, middle temporal gyrus; STG, superior temporal gyrus.

anterior and posterior partition subserving basic syntactic manipulations and semantic associations, respectively.⁵⁶ These functional and structural evidences jointly point to a combinatorial role of the left anterior temporal lobe in compositional semantic operations as well as in the basic syntactic building in the current study.

The right posterior superior temporal cortex was activated likely for the prosodic computation as well as the complex syntactic analysis. On one hand, this area was recruited probably to cope with greater prosodic processes involved in the intelligible speech than in unintelligible backward sounds. 13,57,58 On the other hand, the complex syntax needed to be addressed in the forward but not the backward audios might also contribute to the greater activity of this area in the forward > backward contrast. 8,59

It is known that Chinese sentences do not entail grammatical inflections that mark word category (eg, happy, happiness), case (eg, I wish, he wishes), number (eg, cat, cats), time tense (eg, is, was, been), person (eg, he, him), and so on.^{2,3} Thus, the semantic access of Chinese sentences may require emphasis on the serial word order imposed by syntactic structure, where every word in a sentence must be embedded in an appropriate position and compatible with the words preceding and following it; scrambling the word order will alter the sentence meaning in its entirety. 4-6 Therefore, the syntactic analysis represented in the posterior superior temporal lobe was probably at a relatively complex level, such as sequencing compositional audio words and updating context accordingly, contrasted to the left anterior temporal lobe that preferentially regulated basic syntactic building. The respective parts of the left anterior temporal lobe and posterior superior temporal cortex in basic syntactic building and complex syntactic analysis likely implied that grammatical processing was not localized to one single region, but rather was instantiated in a network involving both the anterior and the posterior temporal systems. 11,60

Despite the observation of expected temporal activations, neural activations outside the temporal cortex were largely out of expectation. First, unpredicted activations were observed in the left sensorimotor area (BA6), which was likely elicited by the covert articulatory repetition more likely occurring in the forward than in the backward audio listening. In addition, several expected regions did not show significant activations in the forward > backward contrast. These regions included the occipitotemporal cortex, the left middle frontal gyrus, and the left ventral inferior frontal cortex.

The lack of activity in the occipitotemporal visual cortex possibly implied that the homophone disambiguation during Chinese sentence listening may require less visualization of heard speech into the written form, owing to the scaffold of sentence context and syntax. This was contrasted to the involvement of orthography in Chinese word listening as indicated in past studies. ^{21–24} Although the left occipitotemporal cortex was reported in Xu et al's study (2013), it could not be taken as evidence for the involvement of orthography in Chinese sentence listening, given the major focus of that study on lexicosemantic rather than sentence-level semantic processing despite the use of sentence stimuli. ²⁷

In fact, the null finding of the left middle frontal gyrus in the forward > backward contrast may also stem from the lower need for speech-to-print conversion in Chinese sentence listening, given the role of the left middle frontal gyrus in assigning the represented syllable to a Chinese character. ^{16,61} On the other hand, the general absence of significant left inferior/middle frontal activations in the forward > backward contrast was more likely attributable to the task effect. ⁶² Listening to normal sentences with no response required might invoke a relatively low cognitive load in the forward condition as in the backward baseline. Thus, subtracting the neural activity of the backward from the forward condition could have cancelled out the expected left frontal activations, given the possible role of the left frontal lobe in executive control. ^{63–65}

In sum, expected activations were observed in the temporal cortex, including the left middle temporal gyrus, the left anterior superior temporal gyrus, and the right posterior superior temporal gyrus. However, expected activation in the occipitotemporal visual cortex was not observed, where visualization of heard sentences to the written forms may be less necessary for homophone discrimination given the availability of contextual scaffolding in sentence processing. Likely due to task sensitivity, expected activations were not shown in the left middle frontal gyrus and the left ventral inferior frontal gyrus, while unexpected activations were noted in the left sensorimotor cortex.

Taken together, the current data substantiated the involvement of the left middle temporal gyrus, the left anterior superior temporal cortex, and the bilateral posterior superior temporal lobes in Chinese auditory sentential semantic neuro-network. This network seems to largely resemble the auditory semantic network for alphabetic languages, with minor specificity noted. Likely tuned to the specific linguistic nature of Chinese, the semantic processing of spoken Chinese sentences appears to elicit little dorsal-stream

activation in the left inferior parietal lobule, while having greater recruitment of the ventral stream via the temporal system.

The Functional Connectivity Pattern Within the Predefined Network

Within the predefined network, none of the ROI-to-ROI functional connectivity was significant in the forward > backward contrast. The null result was likely because the higher-order interregional coordination was less needed by the forward-story passive listening where no responses were required. Meanwhile, the completely incomprehensible reversed sounds in the backward baseline might have involved participants in a semi-resting mental state. This may allow self-thoughts or contemplation to occur, which might induce semantic-like co-activations in the backward baseline, obscuring the functional connectivity patterns that might have been present in the forward speech listening. However, these are speculations which requires further investigations with better task paradigms to more precisely isolate the interconnected network underlying Chinese auditory sentential semantic processing.

Limitations

The current study employed an archival dataset that may lead to some concerns for the validity of the task paradigm for examining auditory sentence comprehension. It is plausible that linguistic components other than semantic processing of sentences differed between the forward and backward task conditions. These aspects of non-interest included syntactic parsing, prosody, word segmentation, lexical access, and temporal sequence, and thus could partially confound activations supposedly observed for sentential semantic processing. However, the forward and the backward sounds were comparable in low-level acoustic stimulation (Table 1), allowing for the semantic and syntactic aspects to be mainly distinguished after the forward > backward comparison. Nevertheless, future studies may still consider using a baseline such as "musical rain," which is also scrambled and unintelligible, but with fine-temporal dynamics more closely matched to the normal speech. With a better matched baseline, the semantic component of interest could be more precisely localized in the task > baseline contrast, and the functional connectivity patterns obscured in the current study are more likely to be clarified. The above limitations may not cover all the possible weaknesses in the current study, but they might provide some worthy points of consideration for future research.

Conclusions

Our data suggest that Chinese auditory sentence comprehension may involve a neural network comprising the left middle temporal gyrus, the left anterior temporal lobe, and the bilateral posterior superior temporal lobes. The occipitotemporal visual cortex was not found to be involved in the sentence-level network of spoken Chinese comprehension, as bottom-up visualization process from homophones to visual forms may be less needed due to the availability of top—down contextual controls in sentence processing. In addition, no significant functional connectivity was observed, likely obscured by the low cognitive demand in the task conditions. This calls for future studies to utilize a more demanding task with a more comparable baseline.

The current Chinese network appears to be generally consistent with the classical networks based on alphabetic languages, but with features specific to Chinese. While the left inferior parietal lobule in the dorsal stream seems to be less involved in the listening comprehension of Chinese sentences, the ventral stream via the temporal cortex appears to be more relevant. These findings deepen our understanding of how the semantic nature of spoken Chinese sentences influences the neural mechanism engaged.

Funding

This work was supported by the Nanyang Technological University – Japan Society for the Promotion of Science (NTU-JSPS) grant and an NTU-SUG grant from Nanyang Technological University. The publication was supported by a research grant (No. BCD1804) from the Bilingual Cognition and Development Lab, National Key Research Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies. The authors would like to thank Dr. Chiao-Yi Wu for providing valuable inputs to improve the manuscript.

Disclosure

The authors report no conflicts of interest in this work.

References

- Li P. Emergent semantic structures and language acquisition: a dynamic perspective. In: Kao Henry SR, Leong CK, & Gao DG, editors. Cognitive Neuroscience Studies of the Chinese Language. Hong Kong: Hong Kong Univ. Press; 2002:79–98.
- Xu Y. Contextual tonal variations in Mandarin. J Phon. 1997;25 (1):61–83. doi:10.1006/jpho.1996.0034
- Ye Z, et al. Semantic and syntactic processing in Chinese sentence comprehension: evidence from event-related potentials. *Brain Res*. 2006;1071(1):186–196. doi:10.1016/j.brainres.2005.11.085

 Comrie B. Language Universals and Linguistic Typology: Syntax and Morphology. University of Chicago press; 1989.

- Perfetti CA, Adlof SM. Reading comprehension: A conceptual framework from word meaning to text meaning. In: Sabatini J, Albro ER, & Oreilly T, editors. *Measuring Up: Advances in How to Assess Reading Ability*. Lanham, MD: Rowman & Littlefield Education; 2012;3:20
- Whaley LJ. Introduction to Typology: The Unity and Diversity of Language. Sage Publications; 1996.
- Friederici AD. The brain basis of language processing: from structure to function. *Physiol Rev.* 2011;91(4):1357–1392. doi:10.1152/ physrev.00006.2011
- 8. Friederici AD. The cortical language circuit: from auditory perception to sentence comprehension. *Trends Cogn Sci.* 2012;16 (5):262–268. doi:10.1016/j.tics.2012.04.001
- Hickok G, Poeppel D. Towards a functional neuroanatomy of speech perception. *Trends Cogn Sci.* 2000;4(4):131–138. doi:10.1016/S1364-6613(00)01463-7
- Hickok G, Poeppel D. The cortical organization of speech processing. Nat Rev Neurosci. 2007;8(5):393–402. doi:10.1038/nrn2113
- 11. Hickok G, Poeppel D. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*. 2004;92(12):67–99. doi:10.1016/j. cognition.2003.10.011
- 12. Price CJ. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*. 2012;62(2):816–847. doi:10.1016/j.neuroimage.2012. 04.062
- McGettigan C, Scott SK. Cortical asymmetries in speech perception: what's wrong, what's right and what's left? *Trends Cogn Sci.* 2012;16 (5):269–276. doi:10.1016/j.tics.2012.04.006
- Cao F, Brennan C, Booth JR. The brain adapts to orthography with experience: evidence from English and Chinese. *Dev Sci.* 2015;18 (5):785–798. doi:10.1111/desc.12245
- Ge J, et al. Cross-language differences in the brain network subserving intelligible speech. *Proc Nat Acad Sci.* 2015;112 (10):2972–2977. doi:10.1073/pnas.1416000112
- Hu W, et al. Developmental dyslexia in Chinese and English populations: dissociating the effect of dyslexia from language differences. *Brain*. 2010;133(Pt 6):1694–1706. doi:10.1093/brain/awq106
- Perfetti CA, Tan L-H. Write to read: the brain's universal reading and writing network. *Trends Cogn Sci.* 2013;17(2):56–57. doi:10.1016/j. tics.2012.12.008
- Schlaggar BL, McCandliss BD. Development of neural systems for reading. Annu Rev Neurosci. 2007;30(p):475–503. doi:10.1146/ annurev.neuro.28.061604.135645
- Tan LH, et al. Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: a meta-analysis. *Hum Brain Mapp*. 2005;25(1):83–91. doi:10.1002/hbm.20134
- Zhu L, et al. Different patterns and development characteristics of processing written logographic characters and alphabetic words: an ALE meta-analysis. *Hum Brain Mapp*. 2014;35(6):2607–2618. doi:10.1002/hbm.22354
- 21. Liu L, et al. Modality-and task-specific brain regions involved in Chinese lexical processing. *J Cogn Neurosci*. 2009;21 (8):1473–1487. doi:10.1162/jocn.2009.21141
- Wu X, Lu J, Chen K, et al. Multiple neural networks supporting a semantic task: an fMRI study using independent component analysis. *Neuroimage*. 2009;45(4):1347–1358. doi:10.1016/j. neuroimage.2008.12.050
- Xiao Z, et al. Differential activity in left inferior frontal gyrus for pseudowords and real words: an event-related fMRI study on auditory lexical decision. *Hum Brain Mapp*. 2005;25(2):212–221. doi:10.1002/hbm.20105
- Zou L, et al. Neural correlates of morphological processing: evidence from Chinese. Front Hum Neurosci. 2016;9.

- Yoncheva YN, et al. Auditory selective attention to speech modulates activity in the visual word form area. *Cerebral Cortex*. 2009;20 (3):622–632. doi:10.1093/cercor/bhp129
- Dehaene S, et al. How learning to read changes the cortical networks for vision and language. Science. 2010;330(6009):1359–1364. doi:10.1126/science.1194140
- Xu G, et al. Access to lexical meaning in pitch-flattened Chinese sentences: an fMRI study. *Neuropsychologia*. 2013;51(3):550–556. doi:10.1016/j.neuropsychologia.2012.12.006
- Han Z, et al. White matter structural connectivity underlying semantic processing: evidence from brain damaged patients. *Brain*. 2013;136(Pt 10):2952–2965. doi:10.1093/brain/awt205
- Li S, et al. Cognitive empathy modulates the processing of pragmatic constraints during sentence comprehension. Soc Cogn Affect Neurosci. 2014;9(8):1166–1174. doi:10.1093/scan/nst091
- Maldjian JA, et al. Multiple reproducibility indices for evaluation of cognitive functional MR imaging paradigms. *Am j Neuroradiol*. 2002:23(6):1030–1037.
- Ashburner J. A fast diffeomorphic image registration algorithm. Neuroimage. 2007;38(1):95–113. doi:10.1016/j.neuroimage.2007.07.007
- Sladky R, et al. Slice-timing effects and their correction in functional MRI. Neuroimage. 2011;58(2):588–594. doi:10.1016/j.neuroimage.2011. 06 078
- Yassa MA, Stark CE. A quantitative evaluation of cross-participant registration techniques for MRI studies of the medial temporal lobe. *Neuroimage*. 2009;44(2):319–327. doi:10.1016/j.neuroimage.2008.09. 016
- Klein A, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage*. 2009;46 (3):786–802. doi:10.1016/j.neuroimage.2008.12.037
- Eklund A, Nichols T, Knutsson H. Can parametric statistical methods be trusted for fMRI based group studies? ArXiv Preprint arXiv. 2015;1511(01863).
- Eklund A, Nichols TE, Knutsson H. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Nat Acad Sci.* 2016;201602413.
- 37. Liu H. Investigating the Behavioral and Neural Correlates of Auditory Semantic Processing in the Chinese Language, in School of Social Sciences. Nanyang Technological University Singapore; 2017:213.
- 38. Brett M, et al. Region of interest analysis using the MarsBar toolbox for SPM 99. *Neuroimage*. 2002;16(2):S497.
- Friston K. Beyond phrenology: what can neuroimaging tell us about distributed circuitry? *Annu Rev Neurosci*. 2002;25(1):221–250. doi:10.1146/annurev.neuro.25.112701.142846
- Stevens MC. The developmental cognitive neuroscience of functional connectivity. *Brain Cogn.* 2009;70(1):1–12. doi:10.1016/j.bandc. 2008.12.009
- Fingelkurts AA, Fingelkurts AA, Kähkönen S. Functional connectivity in the brain is it an elusive concept? *Neurosci Biobehav Rev.* 2005;28(8):827–836. doi:10.1016/j.neubiorev.2004.10.009
- Sala-Llonch R, Bartres-Faz D, Junque C. Reorganization of brain networks in aging: a review of functional connectivity studies. Front Psychol. 2015;6:663. doi:10.3389/fpsyg.2015.00663
- Smith SM, et al. Functional connectomics from resting-state fMRI. Trends Cogn Sci. 2013;17(12):666–682. doi:10.1016/j.tics.2013.09.016
- Whitfield-Gabrieli S, Nieto-Castanon A. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect*. 2012;2(3):125–141. doi:10.1089/brain.2012.0073
- Zalesky A, Fornito A, Bullmore ET. Network-based statistic: identifying differences in brain networks. *Neuroimage*. 2010;53 (4):1197–1207. doi:10.1016/j.neuroimage.2010.06.041
- 46. Walter B, et al. MARINA: An Easy to Use Tool for the Creation of MAsks for Region of INterest Analyses. Paper presented at the 9th International Conference on Functional Mapping of the Human Brain, New York. 2003, June 19–22.

- 47. Humphries C, et al. Time course of semantic processes during sentence comprehension: an fMRI study. *Neuroimage*. 2007;36 (3):924–932. doi:10.1016/j.neuroimage.2007.03.059
- 48. Jobard G, et al. Impact of modality and linguistic complexity during reading and listening tasks. *Neuroimage*. 2007;34(2):784–800. doi:10.1016/j.neuroimage.2006.06.067
- Awad M, et al. A common system for the comprehension and production of narrative speech. *J Neurosci*. 2007;27(43):11455–11464. doi:10.1523/JNEUROSCI.5257-06.2007
- Visser M, Jefferies E, Ralph ML. Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *J Cogn Neurosci*. 2010;22(6):1083–1094. doi:10.1162/ jocn.2009.21309
- Bornkessel I, Schlesewsky M. The extended argument dependency model: a neurocognitive approach to sentence comprehension across languages. *Psychol Rev.* 2006;113(4):787. doi:10.1037/0033-295X.113.4.787
- Brennan J, et al. Syntactic structure building in the anterior temporal lobe during natural story listening. *Brain Lang*. 2012;120 (2):163–173. doi:10.1016/j.bandl.2010.04.002
- Friederici AD, et al. The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cerebral Cortex*. 2003;13(2):170–177. doi:10.1093/cercor/ 13.2.170
- 54. Saur D, et al. Ventral and dorsal pathways for language. *Proc Natl Acad Sci U S A*. 2008;105(46):18035–18040. doi:10.1073/pnas.0805234105
- 55. Wise RJ, Price CJ. Functional neuroimaging of language. In: Cabeza R, Kingstone A, editors. *Handbook of Functional Neuroimaging of Cognition*. Cambridge, MA: MIT Press; 2006:191–228.

- 56. Humphries C, et al. Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Cognitive Neurosci J.* 2006;18(4):665–679. doi:10.1162/jocn.2006.18.4.665
- Xu J, et al. Language in context: emergent features of word, sentence, and narrative comprehension. *Neuroimage*. 2005;25(3):1002–1015. doi:10.1016/j.neuroimage.2004.12.013
- Zatorre RJ, Belin P. Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*. 2001;11(10):946–953. doi:10.10 93/cercor/11.10.946
- Friederici AD, Makuuchi M, Bahlmann J. The role of the posterior superior temporal cortex in sentence comprehension. *Neuroreport*. 2009;20(6):563–568. doi:10.1097/WNR.0b013e3283297dee
- Caplan D, Hildebrandt N, Makris N. Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain*. 1996;119(3):933–949. doi:10.1093/brain/ 119.3.933
- Siok WT, et al. A structural–functional basis for dyslexia in the cortex of Chinese readers. Proc Nat Acad Sci. 2008;105 (14):5561–5566. doi:10.1073/pnas.0801750105
- Cabeza R, Kingstone A. Handbook of Functional Neuroimaging of Cognition. Mit Press; 2006.
- 63. Brass M, et al. The role of the inferior frontal junction area in cognitive control. *Trends Cogn Sci.* 2005;9(7):314–316. doi:10.1016/j.tics.2005.05.001
- 64. Derrfuss J, et al. Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. *Hum Brain Mapp.* 2005;25(1):22–34. doi:10.1002/hbm.20127
- 65. Koechlin E, Ody C, Kouneiher F. The architecture of cognitive control in the human prefrontal cortex. *Science*. 2003;302 (5648):1181–1185. doi:10.1126/science.1088545

Psychology Research and Behavior Management

Publish your work in this journal

Psychology Research and Behavior Management is an international, peer-reviewed, open access journal focusing on the science of psychology and its application in behavior management to develop improved outcomes in the clinical, educational, sports and business arenas. Specific topics covered in the journal include: Neuroscience, memory and decision making; Behavior modification and management; Clinical

applications; Business and sports performance management; Social and developmental studies; Animal studies. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

 $\textbf{Submit your manuscript here:} \ \text{https://www.dovepress.com/psychology-research-and-behavior-management-journal} \\$

