WILEY

RESEARCH ARTICLE

Universal neural basis of structure building evidenced by network modulations emerging from Broca's area: The case of Chinese

Chiao-Yi Wu[1,2] 🄳 | Emiliano Zaccarella[1] | Angela D. Friederici[1]

[1]Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[2]Centre for Research and Development in Learning, Nanyang Technological University, Singapore, Singapore

Correspondence
Chiao-Yi Wu, Centre for Research and Development in Learning, Nanyang Technological University, Academic Building North, 61 Nanyang Drive, ABN-01b-10, Singapore 637335.
Email: cywu@ntu.edu.sg

Abstract

The basic steps in building up language involve binding words of different categories into a hierarchical structure. To what extent these steps are universal or differ across languages is an open issue. Here we examine the neural dynamics of phrase structure building in Chinese—a language that in contrast to other languages heavily depends on contextual semantic information. We used functional magnetic resonance imaging and dynamic causal modeling to identify the relevant brain regions and their dynamic relations. Language stimuli consisted of syntax-driving determiners, semantics-embedded classifiers, and nonverbal symbols making up for two-component sequences manipulated by the factors structure (phrase/list) and number of words (2-word/1-word). Processing phrases compared with word lists elicited greater activation in the anterior part of Broca's area, Brodmann area (BA) 45, and the left posterior superior/middle temporal gyri (pSTG/pMTG), while processing two words against one word led to stronger involvement of the left BA 45, BA 44, and insula. Differential network modulations emerging from subparts of Broca's area revealed that phrasal construction in particular highly modulated the direct connection from BA 44 to left pMTG, suggesting BA 44's primary role in phrase structure building. Conversely, the involvement of BA 45 rather appears sensitive to the reliance on lexico-semantic information in Chinese. Against the background of previous findings from other languages, the present results indicate that phrase structure building has a universal neural basis within the left fronto-temporal network. Most importantly, they provide the first evidence demonstrating that the structure-building network may be modulated by language-specific characteristics.

KEYWORDS

Broca's area, dynamic causal modeling, fMRI, phrases, syntax

1 | INTRODUCTION

The human language, as a highly sophisticated and hierarchical system, comprises elements across a wide range of different scales, from very basic sound segments (i.e., phonemes), basic meaning units (i.e., morphemes), single word components, to phrase constituents and complex sentence structures. The basic steps in building up language involve binding single words of different categories into a hierarchical structure at two levels of combination: one that evaluates and combines the meanings of individual words (i.e., semantic composition), and the other binds the words into a hierarchical structure using abstract rules (i.e., syntactic computation). Syntactic computation is of particular importance as it is considered the basic universal principle that enables the combination of words into hierarchical structures to generate an infinite number of expressions (Chomsky, 1995). Understanding the neural basis of this computation would provide insights into the neurobiological foundation of language—a nature unique to human beings.

To make inferences on the neural mechanisms underlying syntactic computation, previous brain imaging studies have compared

complex sentences with word lists (e.g., Friederici, Meyer, & von Cramon, 2000; Hashimoto & Sakai, 2002; Humphries, Binder, Medler, & Liebenthal, 2006; Indefrey, Hellwig, Herzog, Seitz, & Hagoort, 2004; Matchin, Hammerly, & Lau, 2017; Mazoyer et al., 1993; Pallier, Devauchelle, & Dehaene, 2011; Rogalsky & Hickok, 2009; Snijders et al., 2009; Stowe et al., 1999; Vandenberghe, Nobre, & Price, 2002; Xu, Kemeny, Park, Frattali, & Braun, 2005). Only until recently have more functional magnetic resonance imaging (fMRI) studies utilized minimal phrase structures consisting of two or three words to examine these processes (e.g., Schell, Zaccarella, & Friederici, 2017; Zaccarella & Friederici, 2015a; Zaccarella, Meyer, Makuuchi, & Friederici, 2017). While the studies comparing sentences with word lists have provided valuable understandings of the neural basis of sentence processing, they were inevitably confounded due to the fact that the word lists often comprised a mixture of function and content words (Friederici, 2011). As a result, some basic syntactic combinations were still plausible in the word lists and thus the true effect of syntactic computation may be concealed in the comparison between the sentences and word lists (Zaccarella, Schell, & Friederici, 2017). The studies using minimal phrase structures, on the contrary, have been suggested to eliminate such confounding by minimizing the word lists to two or three words from the same category (Schell et al., 2017; Zaccarella & Friederici, 2015a).

Previous studies that compared two-word determiner phrases with lists of two content words have consistently reported the ventral anterior portion of the left Brodmann Area (BA) 44 to be the neural correlate of syntactic computation in the process of phrase structure building (Schell et al., 2017; Zaccarella & Friederici, 2015a; Zaccarella, Meyer, et al., 2017). A meta-analysis of six studies comparing sentences with lists of only content or function words also supported the involvement of the left BA 44 and the left posterior superior temporal gyrus/sulcus (STG/STS) for syntactic computation (Zaccarella, Schell, et al., 2017). While the left BA 44 has been proposed to be the pure syntactic merger, the left posterior STG/STS plays an integrative role that maps lexico-semantic information to the hierarchy built by the syntactic merger (Den Ouden et al., 2012; Zaccarella, Schell, et al., 2017). Notably, when syntactic sequences were compared with control word-lists that consist of a mixture of content and function words, only the left BA 45 was observed in the left inferior frontal gyrus (Zaccarella, Schell, et al., 2017). It is plausible that the remaining syntactic computations possible in the word lists have canceled out the syntactic effect in the left BA 44. The involvement of the left BA 45, found in the differential neural activation of content and function words, thus is most likely attributed to the processing of semantic information (Bookheimer, 2002; Zaccarella, Meyer, et al., 2017). Taken together, the left fronto-temporal network composed of the left BA 44 and the left posterior temporal cortex, connected via the arcuate fascile/superior longitudinal fascile, are concluded to subserve the neurobiological basis of the basic syntactic computation in human language (Zaccarella & Friederici, 2017; Zaccarella, Schell, et al., 2017).

The left fronto-temporal network for syntactic processing has been supported by neuroimaging evidence from German (e.g., Bornkessel, Zysset, Friederici, von Cramon, & Schlesewsky, 2005; Fiebach, Schlesewsky, Lohmann, von Cramon, & Friederici, 2005; Goucha & Friederici, 2015; Grewe et al., 2005; Wu, Vissiennon, Friederici, &

Brauer, 2016), French (e.g., Pallier et al., 2011; Pattamadilok, Dehaene, & Pallier, 2016), and English (e.g., Constable et al., 2004; Mack, Meltzer-Asscher, Barbieri, & Thompson, 2013). If syntactic computation is the fundamental linguistic capacity in all human languages regardless of typological differences, it is plausible that the same brain region is dedicated to syntactic computation universally. Here in the current study, we examined the neural basis of syntactic computation in Chinese—a language that in contrast to others heavily depends on contextual semantic information.

In an earlier event-related brain potential (ERP) study, Ye, Luo, Friederici, and Zhou (2006) investigated the neural correlates of sentence comprehension by manipulating syntactic violation, semantic violation, or combined syntactic and semantic violations in Chinese ba (把) sentences. The ba structure transforms the Subject–Verb–Object order into the Subject–ba–Object–Verb sequence. The syntactic violation was realized when the verb immediately followed the particle ba with a missing object. The semantic violation occurred when the verb could not be semantically integrated with the prior context built by the subject and the object. In the combined syntactic and semantic violation condition, the verb followed ba immediately without an object and it also could not be integrated into the prior sentential context. They found that combined syntactic and semantic violation processes started in an early time window (at about 150 ms) with the two processes being independent of each other initially (150–250 ms) and interacting at an intermediate phase (250–400 ms). This study reported a sustained negativity around the N400 time window with a stronger effect in the combined violation condition. Other ERP studies (Yu & Zhang, 2008; Zhang, Yu, & Boland, 2010) also observed the semantic-related N400 effect in the combined violation condition, and they concluded that semantic integration proceeded even when syntactic category analysis failed in reading Chinese sentences. These findings are inconsistent with those in German and French which showed that processing sentences with combined syntactic and semantic violations elicited an early left anterior negativity (ELAN) followed by a P600, but no N400 effects. The absence of an N400 effect in the combined violation sentences, which was similar to processing sentences with only syntactic violation, has been suggested to reflect that semantic processing is blocked when syntactic analysis fails—at least in German (Friederici, Gunter, Hahne, & Mauth, 2004; Friederici, Steinhauer, & Frisch, 1999; Hahne & Friederici, 2002) and in French (Isel, Hahne, Maess, & Friederici, 2007). While these ERP studies support a serial, syntax-first model in sentence processing (Frazier & Fodor, 1978; Friederici, 2002), the evidence from Chinese rather suggests a parallel, interactive model implying that in Chinese syntactic category analysis and semantic access are two independent processes and that failed syntactic category analysis does not block semantic access. In light of the differential influences of syntactic and semantic information in the initial stage of sentence processing for different language systems, the current study was set to examine whether syntactic computation in basic phrase structure building has a universal neural basis evidenced also in the case of the Chinese language.

While there is substantial evidence supporting the role of the left fronto-temporal network in syntactic processing, little is known about the information flow within the network. Den Ouden et al. (2012) investigated effective connectivity in the left-lateralized fronto-temporal

network for sentence processing. Crucially, their dynamic causal modeling (DCM) results revealed that the drive to this network during sentence processing originated from the left inferior frontal cortex rather than the posterior superior temporal cortex. The processing of complex sentences modulated the connection from the inferior frontal cortex to the posterior superior temporal cortex, and it was supported by the feedforward and feedback interaction between the inferior frontal cortex for syntactic structure building and the posterior superior temporal cortex for thematic role assignment. The study provides insightful evidence about the driving inputs and information flow within the left fronto-temporal syntactic network. However, the sub-regions within the left inferior frontal gyrus (i.e., BA 44 and BA 45) were not segregated in their investigation, and thus it remains to be elucidated whether the drive of the syntactic network comes from a specific subpart of the frontal region. Another DCM study on reading complex sentences in German reported connectivity from the left BA 44 to the left middle temporal gyrus within the core language system (Makuuchi & Friederici, 2013). In the respective fMRI study (Makuuchi, Grodzinsky, Amunts, Santi, & Friederici, 2013), the activation for processing syntactically complex sentences was found in BA 44, but not in BA 45.

Building on the evidence from the aforementioned studies thus far, we strive to fill in the gap and put the puzzle pieces together to inform the neural mechanism underlying basic phrase structure building at the network level. We utilized fMRI and DCM to examine the functional activation and effective connectivity within the left fronto-temporal language network. The objectives included: (1) to identify which brain regions were involved in phrase structure building in Chinese, (2) to examine whether the drive of phrase structure building came from a segregated subpart within Broca's area, (3) to reveal which connections were modulated by syntactic computation in the process of phrase structure building. Furthermore, the aggregated results from fMRI and DCM analyses would shed light on whether phrase processing in Chinese follows a serial, syntax-first model or a parallel, interactive model. First, our hypothesis for the fMRI analysis (Objective 1) was that phrase structure processing in Chinese would recruit similar cortical regions in the left inferior frontal gyrus (BA 44) and the left posterior temporal cortex as evidenced in previous studies in German (Zaccarella, Schell, et al., 2017). In addition, given that semantic information plays an essential role in Chinese processing, this language-specific characteristic may be reflected in the engagement of semantics-related areas such as the left BA 45 and the left middle temporal gyrus. Second, for the DCM analysis (Objectives 2 and 3), we hypothesized that phrase processing in Chinese would follow a parallel, interactive model, in which driving inputs of phrase structure building would influence both BA 44 and BA 45 in Broca's area, and syntactic computation would then modulate the connections between both regions and the left posterior temporal cortex.

2 | MATERIALS AND METHODS

2.1 | Participants

Since no statistical functional map for the Chinese language on similar tasks as the one used in our experiment (see Task Design and Materials) was available for a power analysis at the time we collected the

data, we determined the sample size on the basis of previous studies from the same laboratory that also investigated the neural underpinnings of phrase processing on simple phrase structures in German. These studies included: Zaccarella and Friederici (2015a), *N* = 22; Zaccarella, Meyer, et al. (2017), *N* = 18; and Schell et al. (2017), *N* = 21. Nineteen healthy native Mandarin Chinese speakers (9 male, 10 female) participated in this study. Prior to the fMRI experiment, all participants passed a medical briefing session conducted by a physician to ensure that they had no history of medical, psychiatric, or neurological disorders. A language background questionnaire (adapted from Marian, Blumenfeld, & Kaushanskaya, 2007) was administered to all interested participants. To reduce potential impact of other languages, if any, on the neural network of sentence processing in Chinese, we only recruited those participants who (1) were native Mandarin Chinese speakers, (2) were highly proficient in Chinese, (3) used Chinese frequently in different settings, and (4) did not take any formal courses to learn other languages before the age of five. One participant was excluded from the analyses due to incidental findings from the MRI scans, resulting in 18 participants (8 males, 10 females; age *M* = 23.9 years, *SD* = 2.10) for the analyses. All were right-handed (*M* = 81.7, *SD* = 23.8) as measured by the Edinburgh Handedness Inventory (Oldfield, 1971). They had no history of language disorders or learning disorders. All participants had Chinese as the primary language of use and reported high proficiency in Chinese. All had learnt English as a second language at an average age of acquisition of 8.89 years (*SD* = 2.56, range 5–13) and almost all of them (17 out of 18) had learnt some German at 20 years of age on average (*SD* = 2.87, range 14–26). The participants' language background profiles (including age of acquisition, ranking of usage, self-reported proficiency, and frequency of usage in different settings for each language) are summarized in the Supporting Information Materials. Overall, through the language background questionnaire, we ensured that our participants had high proficiency and frequency of usage in Chinese. The study was approved by the Institutional Review Board of the University of Leipzig, and informed consent was obtained from all participants.

2.2 | Task design and materials

We investigated basic phrase structure processing in Chinese by using two-word sequences that were either a hierarchical phrase structure or a word list without syntax. We also included one-word sequences to examine the effect of the number of words. Thus, two factors were manipulated in the task, namely the type of Sequence STRUCTURE (phrases or lists) and the number of WORDS (2-word or 1-word). For 2-word conditions, we included sequences in which two Chinese words could either form a simple phrase (i.e., 2-word phrase, 2-PH) or a word list (i.e., 2-word list, 2-LS). Two 1-word conditions—1-word phrase (1-PH) and 1-word list (1-LS)—were created by replacing the second word in the 2-word conditions with a hashtag "#" (Figure 1a). A similar paradigm using German stimuli has been used in Zaccarella and Friederici (2015a).

We employed two-word phrases consisting of a determiner and a classifier (CL) as the primary stimulus structure in the experiment.
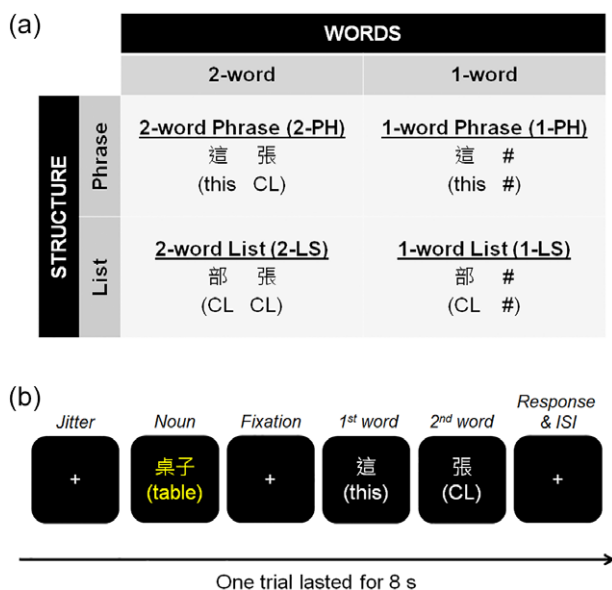
(a)

(b)

One trial lasted for 8 s

FIGURE 1 The experimental design of the sequence judgment task. (a) The task consisted of two levels of the STRUCTURE factor (phrase/list) and two levels of the WORDS factor (2-word/1-word), resulting in four conditions—2-word phrase (2-PH), 1-word phrase (1-PH), 2-word list (2-LS), and 1-word list (1-LS). CL: Classifier. (b) The timing of one trial in the sequence judgment task. Participants were required to judge whether the first and the second words appearing after the fixation cross could be formed into a phrase to describe the noun before. The English translations of the Chinese stimuli were provided in brackets below the Chinese words for illustration, but they were not shown in the actual task. ISI, interstimulus interval [Color figure can be viewed at wileyonlinelibrary.com]

Unlike in German and English, classifiers are obligatory to classify nouns in modern Chinese. Thus, Chinese classifiers are required to be used together with determiners (e.g., demonstratives, quantifiers), forming a classifier phrase to describe nouns. We decided to use classifier phrases (i.e., classifier and determiner) as they consisted of the minimal two-word phrase structure like the determiner phrases (i.e., determiner and noun) in English and German.

Two pilot survey studies were conducted for the selection of the final set of word stimuli. In the first pilot survey, we selected six single-character determiners, including two demonstratives (這 /zhe4/ this, 那 /na4/ that), an article (一 /yi1/ a), and three quantifiers (每 /mei3/ every, 幾 /ji3/ a few, 整 /zheng3/ whole), and eight common classifiers, including 張 /zhang1/, 塊 /kuai4/, 支 /zhi1/, 條 /tiao2/, 根 /gen1/, 座 /zuo4/, 顆 /ke1/, and 隻 /zhi1/. For each classifier, we selected 12 object nouns that could potentially be classified using the particular classifier, resulting in 96 nouns in total. The purpose of the first pilot survey was to assess whether native Chinese speakers would use the same classifier for each noun. Twenty-four native Chinese speakers who did not participate in the fMRI experiment were invited to participate in the first pilot survey. They were provided 96 items in the form of Determiner_Noun (e.g., 這 /zhe4/_桌子 /zhuo1 zi/, this_table) and were asked to fill in the correct classifier for the noun. The determiners were randomly assigned to the nouns. Two classifiers, 支 /zhi1/ and 根 /gen1/, were excluded from the selection process as they were not consistently used for the same

nouns. We selected the nouns for which at least 50% of the participants (n = 12) used the same classifier. As a result, 11 nouns for each of the six classifiers were used for the second pilot survey.

The purpose of the second pilot survey was to assess the acceptability of the phrases selected from the first pilot survey. Fifty native Chinese speakers (29 traditional Chinese users and 21 simplified Chinese users), who did not participate in the first pilot survey nor the fMRI experiment, were invited to rate the acceptability of the noun phrases using a 4-point Likert scale (1—extremely unacceptable, 2—slightly unacceptable, 3—fairly acceptable, 4—fully acceptable). Every noun appeared twice in the survey, one time in a semantically acceptable phrase (e.g., 這張桌子, /zhe4 zhang1 zhuo1 zi/, this CL table), and the other time in a semantically and syntactically unacceptable phrase (e.g., 匹張桌子, /pi3 zhang1 zhuo1 zi/, CL CL table). The latter was created by replacing the determiner with a semantically unrelated classifier, inducing both semantic and syntactic violations in the phrases. The order of phrases was randomized. One participant frequently gave ratings in the opposite direction from the average. Thus, we decided to exclude the participant's response as he/she might have misunderstood the instructions. Based on the other participants' ratings, we excluded the phrases with significantly different ratings between traditional and simplified Chinese users, likely due to different mapping between the classifiers and the nouns in their language conventions. Then we selected 8 nouns for each classifier as the experimental stimuli, and the rest were used in the practice session. The nouns for the experiment were matched in terms of acceptability rating, word length, strokes, and word frequency (according to SUBTLEX-CH; Cai & Brysbaert, 2010) between the classifiers (all $p > .05$).

A total of six determiners, six classifiers, and 48 nouns (eight for each of the six classifiers) comprising 48 correct phrases were included in the experiment. While the mapping between a classifier and a noun was fixed, the pairing of a determiner and a classifier was randomly and equally assigned, so that one determiner occurred with each one of the classifiers for one or two times. In the final list of correct phrases, every determiner and every classifier appeared eight times, and every noun appeared only once. These 48 correct phrases were used as the stimuli for the 2-PH condition. For the stimuli of the 2-LS condition, we replaced the determiners in the correct phrases with another set of classifiers that were not related to the nouns, including 位 /wei4/, 部 /bu4/, 匹 /pi3/, 台 /tai2/, 幅 /fu2/, and 朵 /duo3/. The stimuli for the 1-PH and 1-LS conditions were generated by replacing the second word in the 2-PH and 2-LS conditions with a hashtag #, respectively. Thus, the two levels in the factor STRUCTURE differed in the presence or absence of a phrasal structure, which was realized by the presence or absence of a determiner. Hence, the differences in the word types between the two levels in the factor STRUCTURE were inevitably the manipulation of the phrase/list structures. Each condition consisted of 48 trials. Additionally, 48 null trials with a fixation cross presented at the center of the screen were included.

A total of 240 trials comprising 48 trials for each condition and 48 null trials were pseudorandomly divided into two runs. The following randomization steps were adopted to counterbalance the order of stimuli presented to the subjects. We repeated the pseudorandom division procedure and generated 10 sets, each of which had two runs. Each run consisted of 24 trials for each condition and 24 null

trials, in total 120 trials. We also made sure that the occurrence of each determiner, classifier, and noun was equivalent between the two runs within one set. An event-related design was employed. Within each run, we used an in-house script *conan* to randomize the order of trials and ascertained that no consecutive trials were of the same condition. The 10 sets were repeated with the order of runs reversed, and thus 20 sets of stimuli were generated. Every subject was randomly assigned one set of stimuli, consisting of two runs.

2.3 | Task procedure

Prior to scanning, all participants performed a short practice of the task on a laptop outside the scanner. While determiners and classifiers were the same, none of the nouns used in the practice were repeated in the experimental task. A trial (Figure 1b) started with a fixation cross displayed for a jittered interval of 0, 500, 1,000, or 1,500 ms. A noun was first presented for 600 ms, followed by a fixation cross for 1,800 ms, the first word (determiner or unrelated classifier) for 300 ms, and the second word (related classifier or hashtag) for 300 ms. The trial ended with a fixation cross for a duration of 3,500–5,000 ms, and in total each trial lasted for 8 s. The participants were told that they would see a noun presented in yellow, followed by a cross and then two words in white. They were asked to judge whether the two words after the cross could be combined to form a phrase to describe the noun before the cross. After the second word disappeared, they were required to make a response as quickly and accurately as possible by pressing "1" with their right index finger when the two words could form a phrase, "2" with their right middle finger when the two words could not form a phrase, and "3" with their right ring finger when one of the two words was replaced by a hashtag. In the rest of the time, they were required to look at the fixation cross. Recalling that the same nouns were repeated the same amount of times across all four conditions, our experimental manipulation was specifically designed to test the simple two-word combinatorial effect between the first word (i.e., determiner or classifier) and second word (i.e., classifier or #) in each condition—independently of noun processing, which was orthogonally matched along STRUCTURE (PH vs. LS) and WORDS (2 vs. 1) factors and hence canceled out during functional contrast analysis. In the scanner, all participants underwent two runs of the task. Each run began and ended with a fixation cross presented for 10 and 30 s, respectively. Hence, one run lasted for 16 mins 40 s.

2.4 | Behavioral data analysis

Participants' accuracy and reaction time of the task were computed and examined using two-way repeated measures ANOVA tests.

2.5 | Image acquisition

Scanning was performed in a 3 Tesla MAGNETOM Prisma scanner (Siemens, Erlangen, Germany) with a 32-channel head coil. Functional images were acquired using an echo planar imaging (EPI) sequence with the following parameters: repetition time (TR) 2 s, echo time (TE) 29 ms, flip angle (FA) 77°, field of view 210 mm, matrix size 70×70, in-plane resolution 3×3 mm^2, slice thickness 3 mm, and

37 axial slices acquired bottom-up sequentially with 0.99-mm gaps between slices. A high-resolution 3D MP2RAGE (Magnetization Prepared 2 Rapid Acquisition Gradient Echoes; Marques et al., 2010) sequence (TR 5 s, TE 2.01 ms, FA 4°, matrix size $176 \times 240 \times 256$, resolution $1 \times 1 \times 1$ mm^3) covering the whole brain was used to obtain T1-weighted images as the anatomical reference.

2.6 | Data preprocessing and analysis

Imaging data were preprocessed and analyzed using SPM12 (Wellcome Trust Centre for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm/software/spm12/) under Matlab R2016b (The Mathworks Inc., Natick, MA, USA). All functional images were corrected for slice timing with the middle slice in the acquisition order as the reference, realigned to the first image to correct for head movement, and unwarped with voxel displacement maps generated from field maps to correct for distortions in the magnetic field. The individual's T1-weighted image was coregistered to the mean functional image and then segmented. Functional and anatomical images were normalized to the MNI (Montreal Neurological Institute) space using the DARTEL procedure (Ashburner, 2007), in which a group template was first generated using the participants' gray matter and white matter masks. All the functional and anatomical images were registered to the group template and normalized to the MNI space. Functional images were spatially smoothed with a Gaussian kernel of $7.5 \times 7.5 \times 10$ mm, which was 2.5 times of the nonisotropic voxel size including gaps.

Statistical analyses were performed using a general linear model (GLM). The onsets and duration of trial presentation were convolved with the canonical hemodynamic response function. The incorrect and missed trials were modeled as a separate condition. Therefore, the GLM included six condition regressors: 2-PH, 1-PH, 2-LS, 1-LS, Null, and incorrect/missed trials. Six motion parameters were also added as covariates to account for the variance induced by head movement. Condition-specific contrasts were specified by contrasting each of the four sequence conditions with the Null condition. Subsequently, four condition-specific contrasts were submitted to random-effects group-level analyses. We conducted a two-way flexible factorial analysis with two levels of the STRUCTURE factor (PH, LS) and two levels of the WORDS factor (2-word, 1-word). The main effect of each factor was evaluated to identify the brain regions responsive to the manipulation. Additionally, we examined the effect of conditions (an *F*-test) to reveal the brain regions that showed differential activation to any pair of the conditions. The results of the whole-brain analysis were reported with a cluster-level FWE-corrected threshold of $p < .05$, using the cluster-defining threshold at $p < .001$ (Woo, Krishnan, & Wager, 2014).

2.7 | DCM analysis

Three regions-of-interest (ROIs) were defined in the left BA 44, left BA 45, and left posterior middle temporal gyrus (pMTG). The peak coordinate for each ROI was identified from the effect of conditions in the group-level whole-brain analysis (see Results section) as follows: the left BA 44 (−45, 12, 24), the left BA 45 (−51, 30, 12), and the left pMTG (−42, −54, 20). For individual participants, we specified three volumes-of-interest (VOIs) as 8-mm spheres centering at the

group peak coordinates for each session separately. When the individual participant did not show significant activation at the uncorrected threshold of $p < .05$, the nearest above-threshold voxel was used to create VOIs. Further procedures were taken to ensure that the individual VOIs did not fall out of the respective ROIs. The average distance between the individual peaks from the group peaks were 0.55 mm (SD = 1.48) for the left BA 44, 2.57 mm (SD = 2.49) for the left BA 45, and 3.12 mm (SD = 3.81) for the left pMTG across two sessions.

For DCM analysis, we set up another GLM design matrix at the individual level. In addition to the conditions in the original design matrix, all four task conditions (2-PH, 1-PH, 2-LS, 1-LS) were grouped together as an independent condition (named as Trials), which served as the driving input for DCM. This design matrix, therefore, included seven condition regressors (namely Trials, 2-PH, 1-PH, 2-LS, 1-LS, Null, incorrect-missed trials) and six motion parameters. It was later used to specify the timing of each input in the DCM model specification.

We used DCM12 (Friston, Harrison, & Penny, 2003) implemented in SPM12 to examine effective connectivity in the left fronto-temporal language network for phrase structure building. DCM uses a generic Bayesian framework to make inference about the underlying neuronal states from measured brain activity. It provides posterior estimates of connectivity strength between brain regions and context-dependent modulations on the connections (Stephan et al., 2010). Three types of parameters were estimated in DCM: (1) connectivity strength between brain regions in the absence of external input (i.e., intrinsic connections); (2) changes in the intrinsic connections induced by experimental conditions (i.e., task modulations); (3) direct influence of external input on regional activity (i.e., driving input).

We specified the models with the following parameters: slice timing for each region set as the timing of the reference slice at the slice timing correction in preprocessing (Kiebel, Klöppel, Weiskopf, & Friston, 2007), echo time 0.029 s, bilinear modulatory effect, one state per region, no stochastic effects, and not center input. The model specification was illustrated in Figure 2. The intrinsic connections were defined as bidirectional connections between every two regions for all models. For task modulations, we considered all possibilities that the four conditions could modulate any of the intrinsic connections, which produced $2^6 = 64$ possible models. Three possibilities of driving input were defined, in which the Trials condition could influence only the left BA 44, only the left BA 45, or both the left BA 44 and BA 45. The model space consisted of $64 \times 3 = 192$ models per participant, repeating for 2 sessions. Therefore, 384 models were estimated for each participant.

According to the driving input, we grouped the models with driving input on only the left BA 44 to Family A, those on only the left BA 45 to Family B, and those on both the left BA 44 and BA 45 to Family C. Bayesian model selection was performed on these three families to
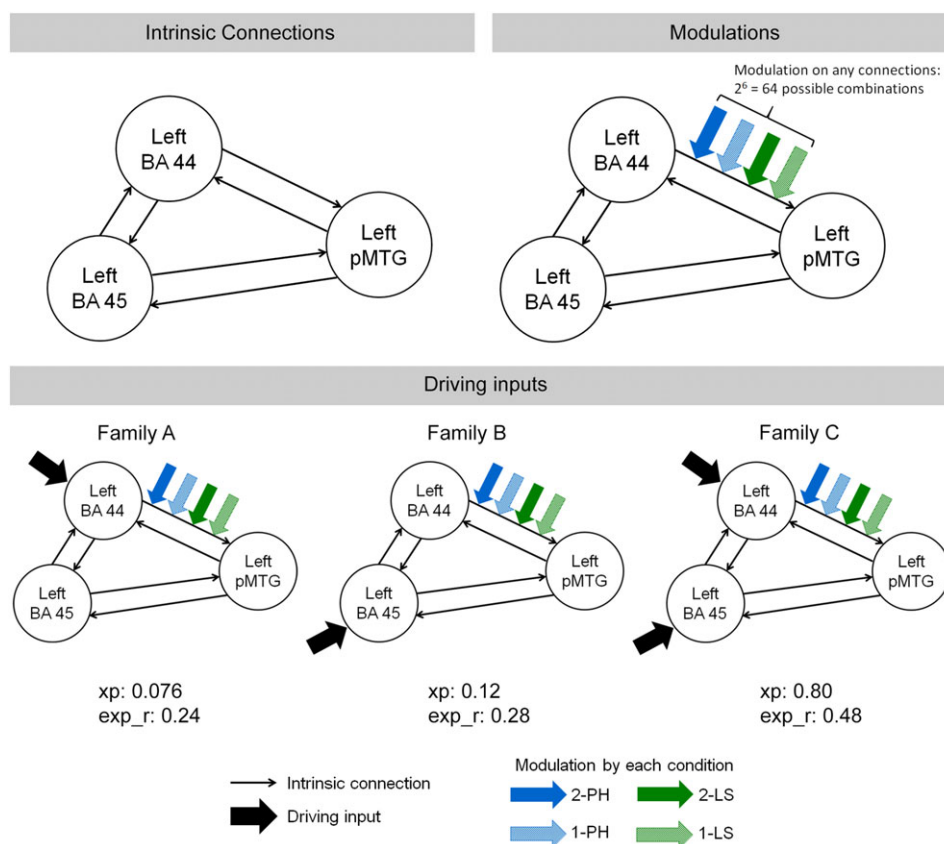

FIGURE 2 Model specification for DCM analysis. All models included bidirectional connections between the regions as the intrinsic connections. We tested all possibilities that the four conditions (2-PH, 1-PH, 2-LS, 1-LS) could modulate any of the six intrinsic connections, resulting in $2^6 = 64$ possible models. Three families of models were defined by the driving inputs, on only the left BA 44 (Family A), on only the left BA 45 (Family B), or on both the left BA 44 and BA 45 (Family C). A total of $2^6 \times 3 = 192$ models, repeating for 2 sessions, were examined. The exceedance probability (xp) and the expected posterior probability (exp_r) of each family are summarized in the figure [Color figure can be viewed at wileyonlinelibrary.com]
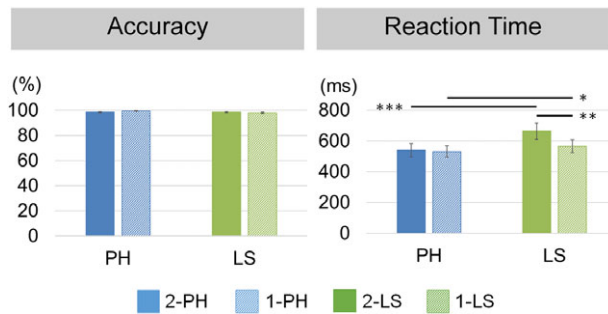
FIGURE 3 Behavioral performance of the sequence judgment task. Both accuracy and reaction time showed a significant main effect of STRUCTURE. A significant interaction between STRUCTURE and WORDS was also found in reaction time. Significant pairwise comparisons are depicted in horizontal lines. 2-PH: 2-word phrase; 1-PH: 1-word phrase; 2-LS: 2-word list; 1-LS: 1-word list. Error bars denote standard errors. * $p < .05$, ** $p < .01$, *** $p < .001$ [Color figure can be viewed at wileyonlinelibrary.com]

determine the most preferred driving input. We then conducted Bayesian model averaging within the winning family to extract the parameter estimates of the intrinsic connections, modulations for each condition, and driving inputs that were weighted by the model evidence of all models within the family (Penny et al., 2010). Posterior probabilities (Pp) were calculated to examine which mean parameters were significantly different from zero (Pp > 0.95), corrected for multiple comparisons using Bonferroni correction.

3 | RESULTS

3.1 | Behavioral results

The behavioral data are illustrated in Figure 3. A main effect of STRUCTURE was found for both accuracy, $F(1,17) = 11.96$, $p = .003$, $\eta_p^2 = .41$, and reaction time, $F(1,17) = 29.1$, $p < .001$, $\eta_p^2 = .63$. The participants had higher accuracy rate for the PH conditions (2-PH: $M = 98.5\%$, $SD = 2.20\%$; 1-PH: $M = 99.6\%$, $SD = 0.97\%$) compared with the LS conditions (2-LS: $M = 98.5\%$, $SD = 2.22\%$; 1-LS: $M = 98.0\%$, $SD = 2.18\%$), and they responded faster to the PH conditions (2-PH: $M = 540$ ms, $SD = 183$; 1-PH: $M = 532$, $SD = 154$) than to the LS conditions (2-LS: $M = 664$, $SD = 224$; 1-LS: $M = 566$, $SD = 176$). In addition, a significant interaction between STRUCTURE and WORDS was found for the reaction time, $F(1,17) = 8.82$, $p = .009$, $\eta_p^2 = .34$. The participants had longer reaction time for the 2-LS condition compared with 2-PH ($p < .001$) and 1-LS ($p = .006$), as well as longer reaction time for the 1-LS condition than the 1-PH condition ($p = .024$). Therefore, the reaction time was included as a covariate in the group-level fMRI analysis.

3.2 | Functional MRI results

The two-way flexible factorial analysis revealed significant activation for the main effect of STRUCTURE and the main effect of WORDS (Table 1 and Figure 4a). For the main effect of STRUCTURE, the PH conditions elicited greater activation in the left middle and superior temporal gyri, the left precentral gyrus, the left pars triangularis (BA 45) of the inferior frontal gyrus, and the left middle frontal gyrus

as compared with the LS conditions. For the main effect of WORDS, increased activation was found for the 2-word conditions as compared with the 1-word conditions in the left pars triangularis (BA 45) and pars opercularis (BA 44) of the inferior frontal gyrus, and the left insula. No significant activation was found in the left fronto-temporal cortex for the interaction effect.

We also examined the effect of conditions, which was used to identify the regions that showed differential activation to any pair of the conditions. The effect of conditions was found in the left postcentral gyrus, the left inferior frontal gyrus, the left angular gyrus and middle temporal gyrus, the right supramarginal gyrus and middle temporal gyrus (Table 1 and Figure 4b). Based on this contrast, we selected the peak activation in the left BA 44 (−45, 12, 24), left BA 45 (−51, 30, 12), and left pMTG (−42, −54, 20) for DCM analysis.

3.3 | DCM results

Bayesian model selection revealed that families A, B, and C had exceedance probabilities of 0.076, 0.12, and 0.80, and expected posterior probabilities of 0.24, 0.28, and 0.48 (Figure 2), respectively. Therefore, family C with driving input on both the left BA 44 and BA 45 was selected as the winning family (Figure 5a). Bayesian model averaging on the parameter estimates of the models in family C further demonstrated that all intrinsic connections and driving inputs were significant (Pp > 0.95, with Bonferroni correction for multiple comparisons). Significant modulation effects were found on different connections for each condition (Figure 5b). The 2-PH significantly modulated the connections from the left BA 44 to left BA 45, from left BA 44 to left pMTG, and from left pMTG to left BA 45. The 1-PH condition had significant modulation effects on the bidirectional connections between the left BA 44 and left pMTG, as well as the connection from the left pMTG to left BA 45. While the 2-LS condition modulated the connections from the left BA 45 to left BA 44 and from the left pMTG to left BA 44, the 1-LS condition only had significant modulation on the connection from the left pMTG to left BA 45. The summary of the parameter estimates of all intrinsic connections, modulations by conditions, and driving inputs is provided in the Supporting Information Table S1. We also performed nonparametric Wilcoxon tests to examine if there were any significant differences in modulations between the conditions. The only significant effect was found in the connection from the left BA 44 to the left pMTG, where the phrase conditions (2-PH: $M = 0.23$, $SD = 0.61$; 1-PH: $M = 0.34$, $SD = 0.65$) had greater modulation effect than the list conditions (2-LS: $M = 0.024$, $SD = 0.73$; 1-LS: $M = 0.052$, $SD = 0.70$), $Z = −2.51$, $p = .012$, with an effect size of $r = .42$ (Pallant, 2007).

4 | DISCUSSION

In the present study, we examined the neural dynamics in the left fronto-temporal language network during phrase structure building in Chinese. First, the whole-brain analysis revealed significant activation in the Broca's area and the left insula for the main effect of WORDS, and in the left BA 45 and the left posterior superior/middle temporal gyri for the main effect of STRUCTURE. Second, the Bayesian model

TABLE 1 Summary of peak coordinates for the whole-brain flexible factorial analysis

Area	MNI coordinates			Cluster size	Z value
	X	Y	Z		
Main effect of STRUCTURE: Phrase > list					
Left middle temporal gyrus	−42	−54	20	136	4.77
Left superior temporal gyrus	−57	−48	20		3.79
Left precentral gyrus	−42	3	44	78	4.27
Left inferior frontal gyrus (BA 45)	−42	39	0	65	4.21
Left middle frontal gyrus	−45	48	0		3.9
Main effect of WORDS: 2-word > 1-word					
Left inferior frontal gyrus (BA 45)	−45	30	8	242	4.62
Left insula	−30	21	0		4.58
Left inferior frontal gyrus (BA 44)	−45	12	24		4.35
Effect of conditions (F-test)					
Left postcentral gyrus	−57	−15	52	296	5.28
Left angular gyrus	−45	−57	24	58	4.53
Left middle temporal gyrus	−45	−63	16		3.48
Left inferior frontal gyrus (BA 45)	−42	39	0	165	4.37
Right supramarginal gyrus	63	−48	40	106	4.22
Right middle temporal gyrus	39	−57	16	78	3.85

The activation clusters were significant at cluster-level FWE-corrected $p < .05$, using the cluster-defining threshold at $p < .001$.

selection procedure in the DCM analysis demonstrated that phrase structure processing in the left fronto-temporal network was driven by linguistic inputs to both the left BA 44 and BA 45 during the sequence judgment task. Third, the Bayesian model averaging analysis showed that differentiated task modulations emerged from the sub-parts of Broca's area, and in particular, the phrase conditions had significantly greater modulations on the connection from the left BA 44 to the left pMTG compared with the list conditions. The results from the whole-brain fMRI and DCM analysis corroborate universal and language-specific aspects for Chinese phrase structure building in the left fronto-temporal network.

4.1 | Neural structures underlying phrase structure building

Our first aim was to identify the neuroanatomical correlates of phrase structure building in Chinese. Both the number of WORDS and the type of STRUCTURE were considered in the analysis. The effect of the number of WORDS indicated that the 2-word conditions elicited greater activation than the 1-word condition in a large cluster encompassing BA 44, BA 45 and the anterior insula in the left hemisphere. As compared with 1-word sequences (i.e., one word followed by a hashtag symbol), reading 2-word sequences engaged increased processing of lexical information regardless of the sequence structure of the stimuli. As several studies have shown that the left inferior frontal gyrus plays a role in phonological and lexico-semantic processing in word recognition (Liuzzi et al., 2017; Vigneau et al., 2006), the greater activation in these areas might be attributed to the increased demands of linguistic processing that were induced by the second word. Notably, the greater activation in the left insula corresponds to the findings in an earlier imaging study that had examined visual word recognition using a similar experimental paradigm in German (Zaccarella &

Friederici, 2015b). The authors reported greater activation in the bilateral insular complexes (greater in the left hemisphere) for reading 2-word as compared with 1-word sequences, with activation peaks being consistently localized in the anterior–dorsal cluster of the insula across participants. The peak of the insular activation in the current study ($x = −30$, $y = 21$, $z = 0$) is in close proximity to their cluster peak at ($x = −33$, $y = 23$, $z = −2$). They (Zaccarella & Friederici, 2017) related the insular activity to word sequencing and described this as a process during which the word is recognized and held for further processing. Therefore, we suggest that the greater activation induced by the 2-word conditions in the left BAs 44/45 and the left insula may be contributed by increased demands of basic recognition processes and further lexical processing, respectively.

The effect of the type of STRUCTURE demonstrated that reading the phrase conditions led to greater activation in the left anterior BA 45 and the left posterior middle and superior temporal gyri as compared with the list conditions. We designed the contrast between the phrase and list conditions in order to reveal the neural correlates of syntactic computation. Our hypothesis was that the basic syntactic computation underlying all languages might have a universal neural representation, but that it remains open whether the neural underpinning might be modulated by language-specific properties as for example in languages that heavily depend on contextual-semantic information. In contrast to findings in German (Zaccarella & Friederici, 2015a), our results of the whole-brain analysis in Chinese did not show a significant main effect of STRUCTURE in the left BA 44 nor an interaction effect between STRUCTURE and WORDS. Instead, the main effect of STRUCTURE was observed in the anterior part of the left BA 45 and the left posterior middle and superior temporal gyri. Note, however, that although left BA 44 did not show significant effects in the whole-brain analysis, the DCM analysis revealed that BA 44 played an important role within the network for phrase
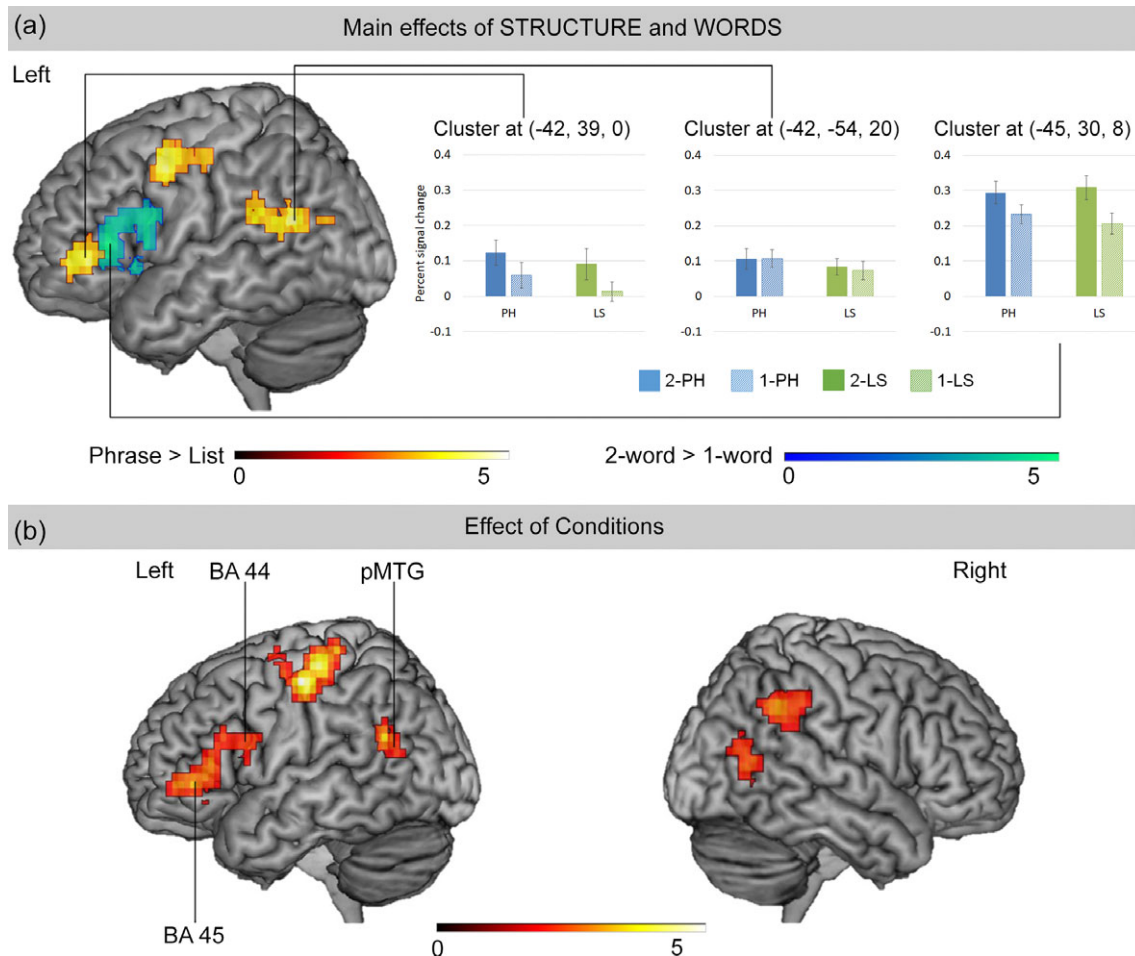
FIGURE 4 The activation maps from the whole-brain flexible factorial analysis for (a) the main effects of STRUCTURE and WORDS, and (b) the effect of conditions (*F*-test). In (a), the activation clusters in warm color showed significant effects of STRUCTURE (greater for the phrase than for the list conditions), and the clusters in cool color showed significant effects of WORDS (greater for the 2-word than for the 1-word conditions). No significant effect was found in the right hemisphere. The bar charts illustrate the percent signal change within the significant clusters for PH > LS in the left inferior frontal gyrus peaked at (−42, 39, 0) and the left middle temporal gyrus peaked at (−42, −54, 20), and for 2-word >1-word in the left inferior frontal gyrus peaked at (−45, 30, 8). In (b), the activation clusters showed significantly differential effects to the comparisons of any two conditions (i.e., effect of conditions). The peak coordinates of the left BA 44, BA 45, and pMTG were used as the center for VOI specification in the DCM analysis. The activation maps were thresholded at cluster-level FWE-corrected *p* < .05, using the cluster-defining threshold at *p* < .001. The color bars represent the range of z-values

structure building (discussed below). Hence, we speculate that the phrase structure building process in Chinese may be overridden by the reliance on lexico-semantic information in the word stimuli, which led to more prominent activation in the left anterior BA 45, rather than BA 44. This result gives rise to the view that the recruitment of the classical language regions is modulated by language-specific characteristics.

4.2 | Neural dynamics: Driving inputs on both left BA 44 and left BA 45

Within the left fronto-temporal language network, BA 44 and BA 45 represent joint driving forces. The Bayesian model selection analysis revealed that the exceedance probability of the family with driving inputs on both the left BAs 44 and 45 outperformed that of the other two families with driving inputs on only the left BA 44 or only the left BA 45. This suggests that the linguistic inputs in Chinese drove phrase structure processing from both the left BAs 44 and 45.

The engagement of both BAs 44 and 45 might indicate that both syntactic and semantic information of the linguistic inputs are processed at the initial stage of word recognition in Chinese. The present results are incongruent with a strong syntax-first model (Frazier & Fodor, 1978; Friederici, 2002). Rather our findings are in line with the previous neurophysiological evidence that suggests a parallel, interactive model for sentence processing in Chinese in which syntactic category analysis and semantic access are two independent processes (Ye et al., 2006; Yu & Zhang, 2008; Zhang et al., 2010). Furthermore, the involvement of the left BA 45 also suggests that semantic information is most relevant in word recognition. In a behavioral study, Wong and Chen (2012) investigated the processing of syntactic category and semantic information in isolated visual word recognition in Chinese. They observed a semantic ambiguity disadvantage in word recognition regardless of whether it was investigated by a lexical decision task, a semantic relatedness judgment task, or a syntactic category judgment task, whereas a syntactic-category ambiguity disadvantage was only observed in the syntactic
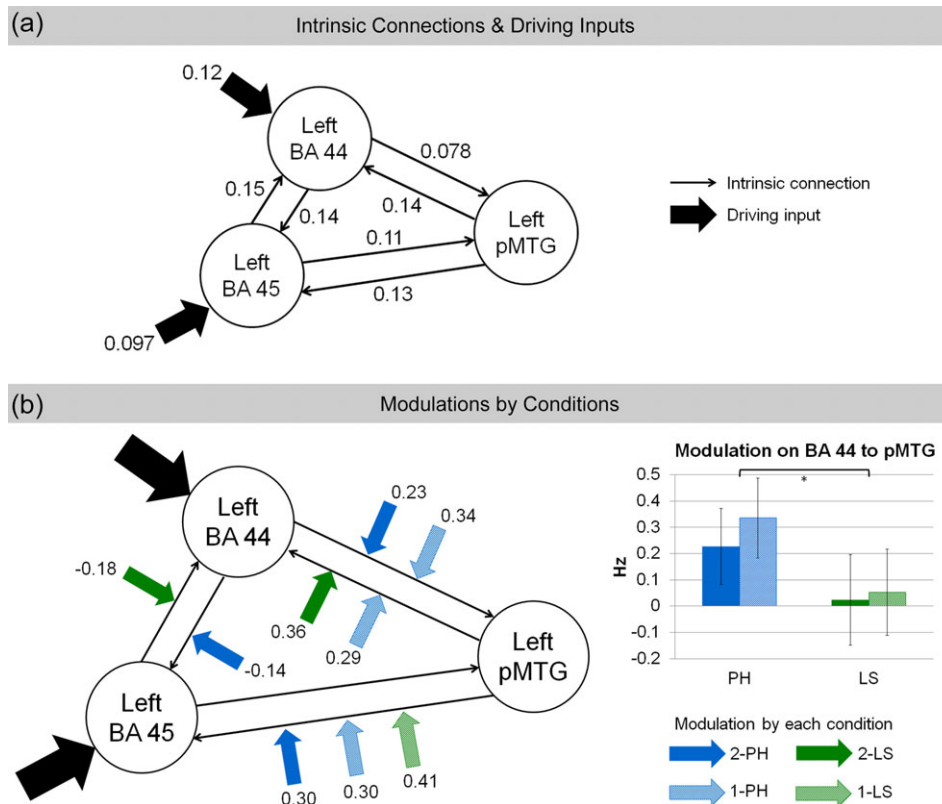
FIGURE 5 The average parameter estimates of the (a) intrinsic connections and driving inputs, and (b) condition modulations of the winning family, which had driving inputs on both the left BA 44 and BA 45. Only significant parameter estimates are shown (posterior probability > 0.95). The phrase conditions (2-PH and 1-PH) had greater modulation effects on the connection from the left BA 44 to left pMTG as compared with the list conditions (2-LS and 1-LS) (*$p < .05$). Black thin arrows: Intrinsic connections; black thick arrows: Driving inputs; blue solid arrows: Modulation from the 2-word phrase condition (2-PH); blue striped arrows: Modulation from the 1-word phrase condition (1-PH); green solid arrows: Modulation from the 2-word list condition (2-LS); green striped arrows: Modulation from the 1-word list condition (1-LS). Error bars denote standard errors [Color figure can be viewed at wileyonlinelibrary.com]

judgment task. Their findings signify the importance of semantic information in Chinese visual word recognition.

4.3 | Modulations within the left fronto-temporal network

Following the initial visual word recognition of the elements in the sequence, each type of sequence structure had differential modulations on the functional connections between the two subparts of Broca's area and the posterior temporal region.

4.3.1 | From frontal to temporal regions

As expected on a universal syntax processing account, phrase structure building even in Chinese involved the left BA 44. Phrase conditions had stronger modulation on the connection from the left BA 44 to the left pMTG as compared with the list conditions. Both 2-PH and 1-PH sequences started with a syntax-driving determiner that opened a hierarchical phrase structure. It has been proposed that the left BA 44 is a syntactic merger and the left posterior superior temporal cortex plays an integrative role in providing lexico-semantic information to the syntactic hierarchy built by the merger (Den Ouden et al., 2012; Zaccarella, Schell, et al., 2017). The stronger modulation of the phrase conditions on the connection from the left BA 44 to the left pMTG corresponds to

the functional roles of these regions as the first determiners open hierarchical phrase structures and await the second words to complete the phrases. Our results show that the functional connection from the left BA 44 to the left pMTG plays a pivotal role in phrase structure building.

It should be noted that our findings are in agreement with previous studies by Den Ouden et al. (2012) and Makuuchi and Friederici (2013). Both studies found that the driving force of complex syntactic processing comes from syntactic computation in the left inferior frontal gyrus, and that complex syntactic structures modulated the connection from the left inferior frontal gyrus to the left posterior temporal cortex. Going beyond these earlier studies, the current study further demonstrates that the modulation effect by linguistic structures can be delineated in the subparts of Broca's area within the left inferior frontal gyrus. Since, however, the current study only examined driving inputs on either the left BA 44 or the left BA 45, or both, further evidence is awaited to understand whether and how additional brain regions beyond Broca's area might provide driving inputs to this left fronto-temporal network for sentence processing.

4.3.2 | Between frontal regions

Our results demonstrated that within Broca's area, the connections between the left BA 44 and the left BA 45 were modulated by

differentiated conditions. The connection from the left BA 44 to the left BA 45 was negatively modulated by the 2-PH condition, whereas the connection from the left BA 45 to the left BA 44 was negatively modulated by the 2-LS condition. Broca's area has been suggested to serve a control and regulation function in language processing to resolve competitions among representations (Novick, Trueswell, & Thompson-Schill, 2010). Along with this interpretation, the left BA 44 and the left BA 45 of Broca's area as well as the connections between them might play a role in controlling syntactic and semantic processing, respectively. For syntactically possible sequences (i.e., 2-PH), the left BA 44 was engaged in syntactic processing and inhibited the aid of semantic information to resolve the phrase structures. On the contrary, for syntactically impossible sequences (i.e., 2-LS) for which no further syntactic processing was required, the left BA 45 came into play to resolve semantic conflicts. The differentiated modulations by the 2-PH and 2-LS conditions on the connections between the left BA 44 and the left BA 45 also imply that the initial processes of syntactic category analysis and semantic access occur in parallel and interact with each other during Chinese phrase processing (Ye et al., 2006; Yu & Zhang, 2008; Zhang et al., 2010).

4.3.3 | From temporal to frontal regions

Feedback connections from the temporal cortex to the frontal cortex also reveal interesting results. The feedback connection from the left pMTG to the left BA 44 was modulated by the 1-PH and 2-LS conditions, both of which included inconsistencies in the formation of phrase structures. The 1-PH condition (Determiner + #) lacked classifiers at the second word position to complete the phrases, while the 2-LS condition (CL + CL) did not have determiners at the first word position to open phrase structures for the second classifiers. The modulation on the connection from the left pMTG to the left BA 44 was significant for these conditions with syntactic inconsistencies but not for the other two conditions which either had correct phrase structures (2-PH condition) or did not open phrase structures at all (1-LS condition). Therefore, this connection might be associated with syntactic repair or detection of syntactic incongruities, where the left pMTG sends feedback information to the left BA 44 when inconsistencies are encountered.

Finally, the connection from the left pMTG to the left BA 45 was significantly modulated by all conditions except for the 2-LS condition. Given the roles of pMTG and the left BA 45 in lexico-semantic processing, this connection might be engaged by word-level and phrase-level semantic evaluation (Hartwigsen et al., 2017).

4.4 | Current limitations and future perspectives

While the present findings put forward the idea of a fundamental linguistic network within which semantic and syntactic information are interactively exchanged between distinct fronto-temporal nodes, the existence of a universal dynamical neurobiological footprint for structure building in the human brain necessarily awaits further evidence to ensure generalizable and language-specific aspects. First, the dynamical model reported here is limited to just one relatively moderate sample size of speakers of one specific language, Chinese, whose mother tongue expertise was strongly controlled in the experimental

setting. Additional languages need to be tested against this dynamical model to ensure functional plausibility. Specifically, second language learners and bilinguals known to show neuroplasticity in these linguistically relevant areas (Kuhl et al., 2016; Mårtensson et al., 2012) might be a crucial testing ground to assess whether the information flow within the network changes according to the degree of proficiency of a certain language, or rather stays independent of proficiency of use and linguistic exposure.

While a clear functional dissociation in the left inferior frontal gyrus between BA 44 and BA 45 with respect to syntax and semantics was demonstrated here, the greater activation reported for the phrase conditions compared with the list conditions in the left precentral gyrus remained less clear. As behavioral judgment was easier for the phrase conditions than it was for the list conditions, it could be possible that the easier conditions yielded more uniform and faster affirmative motor responses, and hence higher activation, than the harder conditions. However, this possibility conflicts with previous evidence showing both increased reaction time and increased activation levels in the area for word category predictions based on syntactic cues without semantic information compared with that with semantic information, even though both conditions required the same affirmative or negative answer (Bonhage, Mueller, Friederici, & Fiebach, 2015).

5 | CONCLUSIONS

To the best of our knowledge, this is the first study to investigate the neural dynamics between the subparts of Broca's area and the left posterior temporal lobe in phrase-level processing by means of fMRI and DCM. Our results revealed differentiated network modulations emerging from BA 44 and BA 45 as subparts of Broca's area in the process of phrase structure building. The connection from the left BA 44 to the left pMTG appears to be crucial for phrase structure building, while the connection from the left pMTG to the left BA 44 provides feedback on syntactic inconsistencies and necessary syntactic repair. The strong involvement of the left BA 45 and its interaction with the other regions in the left fronto-temporal network suggests that in Chinese both semantic and syntactic information are important for phrase structure building. The investigation of the neural representations underlying phrase structure building across different languages thus can shed light on our understanding of the neural basis upon which the human language competence is built (Bolhuis et al., 2014; Ding et al., 2016), and allows to differentiate universal and language-specific aspects of language processing. The current study consolidates the universal roles of the left BA 44 and the left posterior temporal lobe in building syntactic hierarchies, and furthermore demonstrates the language-specific recruitment of the left BA 45 due to higher reliance on semantic information for the Chinese language.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

ORCID

Chiao-Yi Wu https://orcid.org/0000-0001-8237-4038

REFERENCES

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

Bolhuis, J. J., Tattersall, I., Chomsky, N., & Berwick, R. C. (2014). How could language have evolved? *PLOS Biology*, *12*(8), e1001934. https://doi.org/10.1371/journal.pbio.1001934

Bonhage, C. E., Mueller, J. L., Friederici, A. D., & Fiebach, C. J. (2015). Combined eye tracking and fMRI reveals neural basis of linguistic predictions during sentence comprehension. *Cortex*, *68*, 33–47. https://doi.org/10.1016/j.cortex.2015.04.011

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical Organization of Semantic Processing. *Annual Review of Neuroscience*, *25*(1), 151–188. https://doi.org/10.1146/annurev.neuro.25.112701.142946

Bornkessel, I., Zysset, S., Friederici, A. D., von Cramon, D. Y., & Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage*, *26*(1), 221–233. https://doi.org/10.1016/j.neuroimage.2005.01.032

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS One*, *5*(6), e10729. https://doi.org/10.1371/journal.pone.0010729

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: The MIT Press.

Constable, R. T., Pugh, K. R., Berroya, E., Mencl, W. E., Westerveld, M., Ni, W., & Shankweiler, D. (2004). Sentence complexity and input modality effects in sentence comprehension: An fMRI study. *NeuroImage*, *22*(1), 11–21. https://doi.org/10.1016/j.neuroimage.2004.01.001

Den Ouden, D.-B., Saur, D., Mader, W., Schelter, B., Lukic, S., Wali, E., … Thompson, C. K. (2012). Network modulation during complex syntactic processing. *NeuroImage*, *59*(1), 815–823. https://doi.org/10.1016/j.neuroimage.2011.07.057

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158–164. https://doi.org/10.1038/nn.4186

Fiebach, C. J., Schlesewsky, M., Lohmann, G., von Cramon, D. Y., & Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: Syntactic integration versus syntactic working memory. *Human Brain Mapping*, *24*(2), 79–91. https://doi.org/10.1002/hbm.20070

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*(4), 291–325. https://doi.org/10.1016/0010-0277(78)90002-1

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78–84. https://doi.org/10.1016/S1364-6613(00)01839-8

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, *91*(4), 1357–1392. https://doi.org/10.1152/physrev.00006.2011

Friederici, A. D., Gunter, T. C., Hahne, A., & Mauth, K. (2004). The relative timing of syntactic and semantic processes in sentence comprehension. *Neuroreport*, *15*(1), 165–169.

Friederici, A. D., Meyer, M., & von Cramon, D. Y. (2000). Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language*, *74*(2), 289–300. https://doi.org/10.1006/brln.2000.2313

Friederici, A. D., Steinhauer, K., & Frisch, S. (1999). Lexical integration: Sequential effects of syntactic and semantic information. *Memory & Cognition*, *27*(3), 438–453. https://doi.org/10.3758/bf03211539

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, *19*(4), 1273–1302. https://doi.org/10.1016/S1053-8119(03)00202-7

Goucha, T., & Friederici, A. D. (2015). The language skeleton after dissecting meaning: A functional segregation within Broca's area. *NeuroImage*, *114*, 294–302. https://doi.org/10.1016/j.neuroimage.2015.04.011

Grewe, T., Bornkessel, I., Zysset, S., Wiese, R., von Cramon, D. Y., & Schlesewsky, M. (2005). The emergence of the unmarked: A new perspective on the language-specific function of Broca's area. *Human Brain Mapping*, *26*(3), 178–190. https://doi.org/10.1002/hbm.20154

Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research*, *13*(3), 339–356. https://doi.org/10.1016/S0926-6410(01)00127-6

Hartwigsen, G., Henseler, I., Stockert, A., Wawrzyniak, M., Wendt, C., Klingbeil, J., … Saur, D. (2017). Integration demands modulate effective connectivity in a fronto-temporal network for contextual sentence integration. *NeuroImage*, *147*, 812–824. https://doi.org/10.1016/j.neuroimage.2016.08.026

Hashimoto, R., & Sakai, K. L. (2002). Specialization in the left prefrontal cortex for sentence comprehension. *Neuron*, *35*(3), 589–597. https://doi.org/10.1016/S0896-6273(02)00788-2

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience*, *18*(4), 665–679. https://doi.org/10.1162/jocn.2006.18.4.665

Indefrey, P., Hellwig, F., Herzog, H., Seitz, R. J., & Hagoort, P. (2004). Neural responses to the production and comprehension of syntax in identical utterances. *Brain and Language*, *89*(2), 312–319. https://doi.org/10.1016/S0093-934X(03)00352-3

Isel, F., Hahne, A., Maess, B., & Friederici, A. D. (2007). Neurodynamics of sentence interpretation: ERP evidence from French. *Biological Psychology*, *74*(3), 337–346. https://doi.org/10.1016/j.biopsycho.2006.09.003

Kiebel, S. J., Klöppel, S., Weiskopf, N., & Friston, K. J. (2007). Dynamic causal modeling: A generative model of slice timing in fMRI. *NeuroImage*, *34*(4), 1487–1496. https://doi.org/10.1016/j.neuroimage.2006.10.026

Kuhl, P. K., Stevenson, J., Corrigan, N. M., van den Bosch, J. J. F., Can, D. D., & Richards, T. (2016). Neuroimaging of the bilingual brain: Structural brain correlates of listening and speaking in a second language. *Brain and Language*, *162*, 1–9. https://doi.org/10.1016/j.bandl.2016.07.004

Liuzzi, A. G., Bruffaerts, R., Peeters, R., Adamczuk, K., Keuleers, E., De Deyne, S., … Vandenberghe, R. (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

Mack, J., Meltzer-Asscher, A., Barbieri, E., & Thompson, C. (2013). Neural correlates of processing passive sentences. *Brain Sciences*, *3*(3), 1198–1214. https://doi.org/10.3390/brainsci3031198

Makuuchi, M., & Friederici, A. D. (2013). Hierarchical functional connectivity between the core language system and the working memory system. *Cortex*, *49*(9), 2416–2423. https://doi.org/10.1016/j.cortex.2013.01.007

Makuuchi, M., Grodzinsky, Y., Amunts, K., Santi, A., & Friederici, A. D. (2013). Processing noncanonical sentences in Broca's region: Reflections of movement distance and type. *Cerebral Cortex*, *23*(3), 694–702. https://doi.org/10.1093/cercor/bhs058

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Marques, J. P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.-F., & Gruetter, R. (2010). MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *NeuroImage*, *49*(2), 1271–1281. https://doi.org/10.1016/j.neuroimage.2009.10.002

Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., & Lövdén, M. (2012). Growth of language-related brain areas after foreign language learning. *NeuroImage*, *63*(1), 240–244. https://doi.org/10.1016/j.neuroimage.2012.06.043

Matchin, W., Hammerly, C., & Lau, E. (2017). The role of the IFG and pSTS in syntactic prediction: Evidence from a parametric study of

hierarchical structure in fMRI. *Cortex*, *88*, 106–123. https://doi.org/10.1016/j.cortex.2016.12.010

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., ... Mehler, J. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience*, *5*(4), 467–479. https://doi.org/10.1162/jocn.1993.5.4.467

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2010). Broca's area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass*, *4*(10), 906–924. https://doi.org/10.1111/j.1749-818X.2010.00244.x

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.

Pallant, J. (2007). *SPSS survival manual—A step by step guide to data analysis using SPSS for windows* (3rd ed.). Maidenhead: Open University Press.

Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(6), 2522–2527. https://doi.org/10.1073/pnas.1018711108

Pattamadilok, C., Dehaene, S., & Pallier, C. (2016). A role for left inferior frontal and posterior superior temporal cortex in extracting a syntactic tree from a sentence. *Cortex*, *75*, 44–55. https://doi.org/10.1016/j.cortex.2015.11.012

Penny, W. D., Stephan, K. E., Daunizeau, J., Rosa, M. J., Friston, K. J., Schofield, T. M., & Leff, A. P. (2010). Comparing families of dynamic causal models. *PLoS Computational Biology*, *6*(3), e1000709. https://doi.org/10.1371/journal.pcbi.1000709

Rogalsky, C., & Hickok, G. (2009). Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cerebral Cortex*, *19*(4), 786–796. https://doi.org/10.1093/cercor/bhn126

Schell, M., Zaccarella, E., & Friederici, A. D. (2017). Differential cortical contribution of syntax and semantics: An fMRI study on two-word phrasal processing. *Cortex*, *96*, 105–120. https://doi.org/10.1016/j.cortex.2017.09.002

Snijders, T. M., Vosse, T., Kempen, G., Van Berkum, J. J. A., Petersson, K. M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: An fMRI study using word-category ambiguity. *Cerebral Cortex*, *19*(7), 1493–1503. https://doi.org/10.1093/cercor/bhn187

Stephan, K. E., Penny, W. D., Moran, R. J., den Ouden, H. E. M., Daunizeau, J., & Friston, K. J. (2010). Ten simple rules for dynamic causal modeling. *NeuroImage*, *49*(4), 3099–3109. https://doi.org/10.1016/j.neuroimage.2009.11.015

Stowe, L. A., Paans, A. M. J., Wijers, A. A., Zwarts, F., Mulder, G., & Vaalburg, W. (1999). Sentence comprehension and word repetition: A positron emission tomography investigation. *Psychophysiology*, *36*(6), 786–801. https://doi.org/10.1111/1469-8986.3660786

Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience*, *14*(4), 550–560. https://doi.org/10.1162/08989290260045800

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., ... Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*, *30*(4), 1414–1432. https://doi.org/10.1016/j.neuroimage.2005.11.002

Wong, A. W.-K., & Chen, H.-C. (2012). Is syntactic-category processing obligatory in visual word recognition? Evidence from Chinese. *Language and Cognitive Processes*, *27*(9), 1334–1360. https://doi.org/10.1080/01690965.2011.603931

Woo, C.-W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *NeuroImage*, *91*, 412–419. https://doi.org/10.1016/j.neuroimage.2013.12.058

Wu, C.-Y., Vissiennon, K., Friederici, A. D., & Brauer, J. (2016). Preschoolers' brains rely on semantic cues prior to the mastery of syntax during sentence comprehension. *NeuroImage*, *126*, 256–266. https://doi.org/10.1016/j.neuroimage.2015.10.036

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *NeuroImage*, *25*(3), 1002–1015. https://doi.org/10.1016/j.neuroimage.2004.12.013

Ye, Z., Luo, Y.-j., Friederici, A. D., & Zhou, X. (2006). Semantic and syntactic processing in Chinese sentence comprehension: Evidence from event-related potentials. *Brain Research*, *1071*(1), 186–196. https://doi.org/10.1016/j.brainres.2005.11.085

Yu, J., & Zhang, Y. (2008). When Chinese semantics meets failed syntax. *Neuroreport*, *19*(7), 745–749. https://doi.org/10.1097/WNR.0b013e3282fda21d

Zaccarella, E., & Friederici, A. D. (2015a). Merge in the human brain: A subregion based functional investigation in the left pars Opercularis. *Frontiers in Psychology*, *6*, 1818. https://doi.org/10.3389/fpsyg.2015.01818

Zaccarella, E., & Friederici, A. D. (2015b). Reflections of word processing in the insular cortex: A sub-regional parcellation based functional assessment. *Brain and Language*, *142*, 1–7. https://doi.org/10.1016/j.bandl.2014.12.006

Zaccarella, E., & Friederici, A. D. (2017). The neurobiological nature of syntactic hierarchies. *Neuroscience & Biobehavioral Reviews*, *81*, 205–212. https://doi.org/10.1016/j.neubiorev.2016.07.038

Zaccarella, E., Meyer, L., Makuuchi, M., & Friederici, A. D. (2017). Building by syntax: The neural basis of minimal linguistic structures. *Cerebral Cortex*, *27*(1), 411–421. https://doi.org/10.1093/cercor/bhv234

Zaccarella, E., Schell, M., & Friederici, A. D. (2017). Reviewing the functional basis of the syntactic merge mechanism for language: A coordinate-based activation likelihood estimation meta-analysis. *Neuroscience & Biobehavioral Reviews*, *80*, 646–656. https://doi.org/10.1016/j.neubiorev.2017.06.011

Zhang, Y., Yu, J., & Boland, J. E. (2010). Semantics does not need a processing license from syntax in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 765–781. https://doi.org/10.1037/a0019254

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.