

# Syntactic processing of Mandarin Chinese as a second language recruits a crucial frontoparietal network

## Research Article

**Cite this article:** Gao, C., Guo, J., Zhang, L., Li, Z., Chen, L. and Feng, L. (2025). Syntactic processing of Mandarin Chinese as a second language recruits a crucial frontoparietal network. *Bilingualism: Language and Cognition* 1–18. <https://doi.org/10.1017/S1366728925100138>

Received: 14 January 2024

Revised: 4 May 2025

Accepted: 13 May 2025

### Keywords:

Chinese as a second language; function word; hierarchical syntactic processing; fMRI; learner factors

### Corresponding authors:

Luyao Chen and Liping Feng;  
Emails: [luyaochen@bnu.edu.cn](mailto:luyaochen@bnu.edu.cn);  
[fengliping@bnu.edu.cn](mailto:fengliping@bnu.edu.cn)

C.G. and J.G. have contributed equally to this work and shared first authorship.

Chenyang Gao<sup>1,2</sup> , Jia Guo<sup>3,4</sup>, Liaoyuan Zhang<sup>5</sup>, Zimu Li<sup>2</sup>, Luyao Chen<sup>6,7,8</sup>  and Liping Feng<sup>2</sup>

<sup>1</sup>School of Global Education and Development, University of Chinese Academy of Social Sciences, Beijing, China; <sup>2</sup>School of International Chinese Language Education, Beijing Normal University, Beijing, China; <sup>3</sup>School of Chinese Language and Literature, Beijing Foreign Studies University, Beijing, China; <sup>4</sup>State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China; <sup>5</sup>Faculty of Humanities and Foreign Language Education, Beijing Institute of Education, Beijing, China; <sup>6</sup>Max Planck Partner Group, School of International Chinese Language Education, Beijing Normal University, Beijing, China; <sup>7</sup>Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany and <sup>8</sup>Institute of Educational System Science, School of Systems Science, Beijing Normal University, Beijing, China

### Abstract

Previous L1 syntactic processing studies have identified the crucial left frontotemporal network, whereas research on L2 syntactic processing has shown that learner factors, such as L2 proficiency and linguistic distance, can modulate the related networks. Here, we developed a function-word-based jabberwocky sentence reading paradigm to investigate the neural correlates underlying Chinese L2 syntactic processing. Twenty Chinese L2 Korean native speakers were recruited in this fMRI study. Chinese proficiency test scores and Chinese-Korean syntactic similarity scores were measured to quantify the learner factors, respectively. The imaging results revealed an effective left frontoparietal network involving superior parietal lobule (SPL), posterior inferior frontal gyrus (pIFG) and precentral gyrus (PreCG). Moreover, the signal intensity of SPL as well as the connectivity strength between SPL and PreCG significantly correlated with the learner factors. These findings shed light on the neurobiological relationships between L1 and L2 syntactic processing and on the modulation of L2 learner factors.

## Highlights

- The left SPL is crucial in Chinese L2 syntactic processing.
- Chinese L2 syntactic processing relies on the pivotal frontoparietal network.
- The linguistic distance and L2 proficiency can modulate the frontoparietal network.
- Current findings support the Shallow Structure Hypothesis.

## 1. Introduction

The faculty of language (FL) in the narrow sense relies on Merge, a basic syntactic computation unique to the human species (Hauser et al., 2002). Briefly speaking, Merge is the process of combining two syntactic objects (e.g., “the” and “apple”) into a larger structure (i.e., “the apple” [a determiner phrase]) (Chomsky, 1995). The realization of Merge requires the identification/labeling of the syntactic categories of the syntactic objects as well as the merged constituents. Thus, the processing of complex syntactic structures through multi-level Merge can reflect the hierarchical nature of the human language faculty (Hauser et al., 2002).

How about Merge in the second language (L2) processing? Clahsen and Felser (2006a, 2006b, 2006c) proposed the *Shallow Structure Hypothesis* (SSH), which suggested that the syntactic representations of L2 speakers are not as deep and detailed as those of native speakers (L1) but rather shallower and more ambiguous. We chose the SSH as one of the prominent hypotheses to begin with because it was specifically proposed based on the difficulties faced by L2 learners in online processing of ambiguous relative clauses. It pointed out that the automation level of complex syntactic processing in L2 learners was limited, relying on shallow information, such as lexical, semantic, discourse and pragmatic cues, which fundamentally differed from the processing mechanisms of native speakers. Researchers found that even high-level L2 learners still have difficulty in constructing abstract syntactic representations in real time, both at the sentence level (e.g., relative-clause ambiguities and filler-gap dependencies, Clahsen & Felser, 2006a, 2006c) and at the microscopic level of morphological features (e.g., inflectional and derivative affixes, Kirkici & Clahsen, 2013). These studies mostly used natural materials to explore whether L2 learners underuse syntactic information. However, processing differences between L1 and L2 often

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use and/or adaptation of the article.

showed up only where the manipulation of abstract grammatical representations was required (Clahsen & Felser, 2018). Therefore, it remains to be specified whether the core syntactic computation of hierarchical structures (i.e., Merge) could be predicted by SSH during the pure L2 syntactic processing as well.

Moreover, SSH has extended itself to the level of the time course differences of processing different information types. Studies employing eye movements or event-related potentials (ERP) techniques (see review in Clahsen & Felser, 2018) found that the processing strategies for syntactic information in L2 might differ from those in L1. For instance, Guo et al. (2009) observed that Chinese-native English L2 learners elicited N400, rather than P600 in English native speakers, when processing sentences with verb sub-categorization violations. This suggested that native speakers primarily relied on syntactic information, whereas L2 learners depended more on semantic information. Similarly, Zhang's (2019) ERP study pointed out that, unlike L1, Chinese L2 learners did not report P600 after syntactically violated verbs, reflecting L2's greater reliance on lexical and pragmatic knowledge. In contrast, a stream of studies has shown that advanced L2 learners could exhibit ERP patterns similar to those of native speakers in syntactic violation paradigms (Bowden et al., 2013; Rossi et al., 2006; Steinhauer et al., 2009). For example, Steinhauer et al. (2009) reported that highly proficient French and Chinese native speakers learning English exhibited a biphasic LAN/P600 pattern consistent with that of native English speakers when processing syntactic category violations. These two ERP components were interpreted as reflecting the early disruption of relatively automated syntactic processing (LAN) and subsequent reanalysis and repair of sentence structure (P600). In line with these studies, Chen et al. (2023a) also identified a similar ERP pattern of the high-level Chinese L2 learners with that of the Chinese native speakers. Moreover, in artificial grammar paradigms involving (morpho) syntactic violations, highly proficient learners also demonstrated a (LAN-) P600 pattern (Friederici et al., 2002; Grey et al., 2018; Morgan-Short et al., 2010, 2012a, 2012b). Hence, it is still disputed to what extent the time-course differences might or might not support SSH. More crucially, differences in the spatial dimension, that is, the brain activation differences between L1 and L2 syntactic processing, are largely unspecified.

Although SSH lacks a specific assumption in the spatial dimension, existing bilingual studies have reported that even though L1 and L2 share activated brain regions partially, L2 processing still maintains its own activation patterns (Scherer et al., 2012; Suh et al., 2007). Suh et al. (2007) reported significant L1–L2 overlapping activations of the left IFG, the bilateral inferior parietal and occipital lobes. However, in L1, embedded sentences resulted in stronger activation of the IFG compared to conjoined sentences, whereas no such difference was observed in L2. Similarly, Scherer et al. (2012) found significant L1–L2 overlapping activation in left frontal regions during syntactic processing, suggesting a common pattern of brain activity in highly proficient bilinguals. While the study further revealed that L2 recruited more temporal areas but fewer frontal areas compared to L1, suggesting L2 processing was more demanding and less automatic.

Such differences might be attributed to the L2 learner factors, as previous studies have concluded that the experience of learning and using L2 can have a pervasive impact on the structure and function of the cerebral cortex (Kovelman et al., 2008; Mechelli et al., 2004). For example, functionally, compared with native speakers, L2 learners increased the signal intensity of the left IFG during an L2 grammatical judgment task (Kovelman et al., 2008). And these

effects were moderated primarily by learner factors such as native language backgrounds and L2 proficiency. A recent meta-analysis by Cargnelutti et al. (2022) indicated that L2 proficiency and linguistic distance between L1 and L2 worked together in L2-related neural networks, but the researchers also pointed out that empirical studies quantifying linguistic distance were relatively lacking. Given the variability in individual L2 proficiency and the differences in proficiency across structures influenced by linguistic distance (Steinhauer et al., 2009), researchers have emphasized the need to account for individual variation when investigating L2 processing. They suggested treating individual difference variables as continuous rather than categorical variables to enable a more nuanced and detailed understanding of the L2 factors (Hell, 2023). Therefore, how learner factors would modulate the neural basis of L2 syntactic processing still awaited to be specified.

It is noteworthy that SSH as well as the related aforementioned studies on L2 syntactic processing concentrated on the morphosyntactic information within the scope of Indo-European languages. As for Chinese, an isolated language lacking morphosyntactic information, function words play a critical role in conveying syntactic relations as a distinct feature of Chinese syntax (Huang & Liao, 2011; Tang, 2019). Based on the function words, we can better understand the neurobiology of L2 syntactic processing with prominent Chinese features.

Previous investigations on Chinese L2 syntactic processing have predominantly adopted natural language materials (see Zhao, 2018 for a review), which are prone to interference from semantic processing. As complex structures can be constructed based on abstract syntactic information (such as inflectional morphemes and function words), several syntactic processing studies adopted the *jabberwocky* paradigm (Matchin et al., 2017; Ohta et al., 2013; Pallier et al., 2011), replacing content words with pseudowords, which is able to not only better control the interference of the confounding factors but also to investigate the natural syntactic rules in an ecological way by keeping the real syntactic structures intact (see a recent review by Maran et al., 2022). Inspired by these studies, a previous study developed a Chinese function-word-based syntactic processing paradigm deprived of conceptual-semantic information (Chen et al., 2023b). The experimental materials of complex structures were generated by real Chinese function words and pseudo-content-words following real Chinese syntactic rules, and participants were asked to label the *jabberwocky* sequences with the corresponding syntactic categories (e.g., noun and verb phrases). They found that Chinese L1 speakers recruited a language-general frontotemporal syntactic network dominant in the left hemisphere, including both the posterior IFG (esp., Broca's area) as well as the posterior temporal lobe (pTL), consistent with previous findings conducted on inflecting language (e.g., Matchin et al., 2017; Pallier et al., 2011; Zaccarella et al., 2017). Therefore, whether and to what extent Chinese L2 learners share the same neural basis as L1 speakers during syntactic processing based on Chinese function words to evidence SSH should be substantially scrutinized.

Thus, we adopted the well-established Chinese function-word-based *jabberwocky* sentence reading paradigm and recruited relatively proficient Korean native speakers to explore the neurobiology of Chinese L2 syntactic processing. The participants needed to identify the syntactic category of the holistic structure through merging the syntactic elements according to the function words. Besides, Chinese L2 learner factors, such as linguistic distance and L2 proficiency, were measured as continuous variables.

To sum up, our research questions are as follows:

- (1) What is the neural basis of Chinese L2 hierarchical syntactic processing based on the function words?
- (2) Will and how the learner factors (esp., L2 proficiency and linguistic distance) modulate Chinese L2 syntactic processing at the neurobiological level?

We expected that: Chinese L2 learners may share core brain regions such as the left posterior IFG with native speakers during syntactic processing, but the activated brain regions and connectivity networks were modulated by L2 learner factors so that differed from the frontotemporal network.

## 2. Materials and methods

### 2.1. Participants

Twenty native Korean speakers at advanced Chinese level were recruited (male: 10, age:  $M = 25.55$  years,  $SD = 2.48$  years). They started learning Chinese at about  $16.40 \pm 4.66$  years old and were classified as late L2 learners. Rationales of the recruitment of Korean native speakers were explained in Section 1.1 of the Supplementary Material.

Advanced Chinese L2 learners were chosen because they already had basic knowledge of Chinese function words, which enabled us to better investigate the process of syntactic processing rather than learning. The screening criteria were as follows: (1) Participants who had passed the HSK Level 5 and Level 6 from the Korean students studying in Beijing; (2) Using the Chinese L2 proficiency rapid test paper compiled by Feng et al. (2020). Participants with scores above 18 were included in the experiment (full score is 30). The average score in this experiment is  $26.6 \pm 2.97$  points, and each participant's score would be used as a quantitative indicator for evaluating his/her Chinese L2 proficiency.

All participants were right-handed, had normal or corrected-to-normal vision, and had no visual impairments, no dyslexia, and no history of psychiatric or neurological diseases. Because the visual materials included Japanese Katakana and the Chinese Phonetic Alphabet (i.e., the Taiwanese Bopomofo, Zhuyin), none of the participants had experience in learning Japanese and the Chinese Phonetic alphabet, and they did not major in linguistics, psychology or neuroscience. All participants signed the "Experiment Informed Consent" and received remuneration afterwards. Participants' data were analyzed according to the criteria of head motion artifacts ( $< 2$  mm in translation and  $< 2^\circ$  in rotation). This study was approved by the local ethics committee of Beijing Normal University, Beijing, China.

### 2.2. Materials

The details of the experimental materials and procedures of this study were consistent with the previous Chinese L1 study (Chen et al., 2023b), except for the part involving L2 learners as well as the L2 learner factors. For a brief description of the stimuli and procedures, see also below.

The materials consisted of Chinese function words and pseudo-content-words (Figure 1A). These function words are highly abstract without concrete meanings (Huang & Liao, 2011; Tang, 2019), but they can combine with other pseudo-content-words (Figure 1A) to generate well-formed structures (Figure 1B). According to the "Chinese Proficiency Grading Standards for International Chinese Language Education" (Ministry of Education of the People's

Republic of China, 2021), we controlled the difficulty of these function words at the elementary to intermediate level to reduce vocabulary difficulties.

Pseudo-content-words were generated by using unfamiliar Chinese-character-like symbols, including Japanese Katakana and Taiwanese Bopomofo (Zhuyin), which were similar to Chinese orthography, but unpronounceable and semantic-free for the participants (Figure 1A). These symbols were arbitrarily assigned syntactic categories. The semantic strategy to relate the pseudo-words to the real words for each pseudo-content-word was also controlled (see Section 1.2 of the Supplementary Material and Figure 1A).

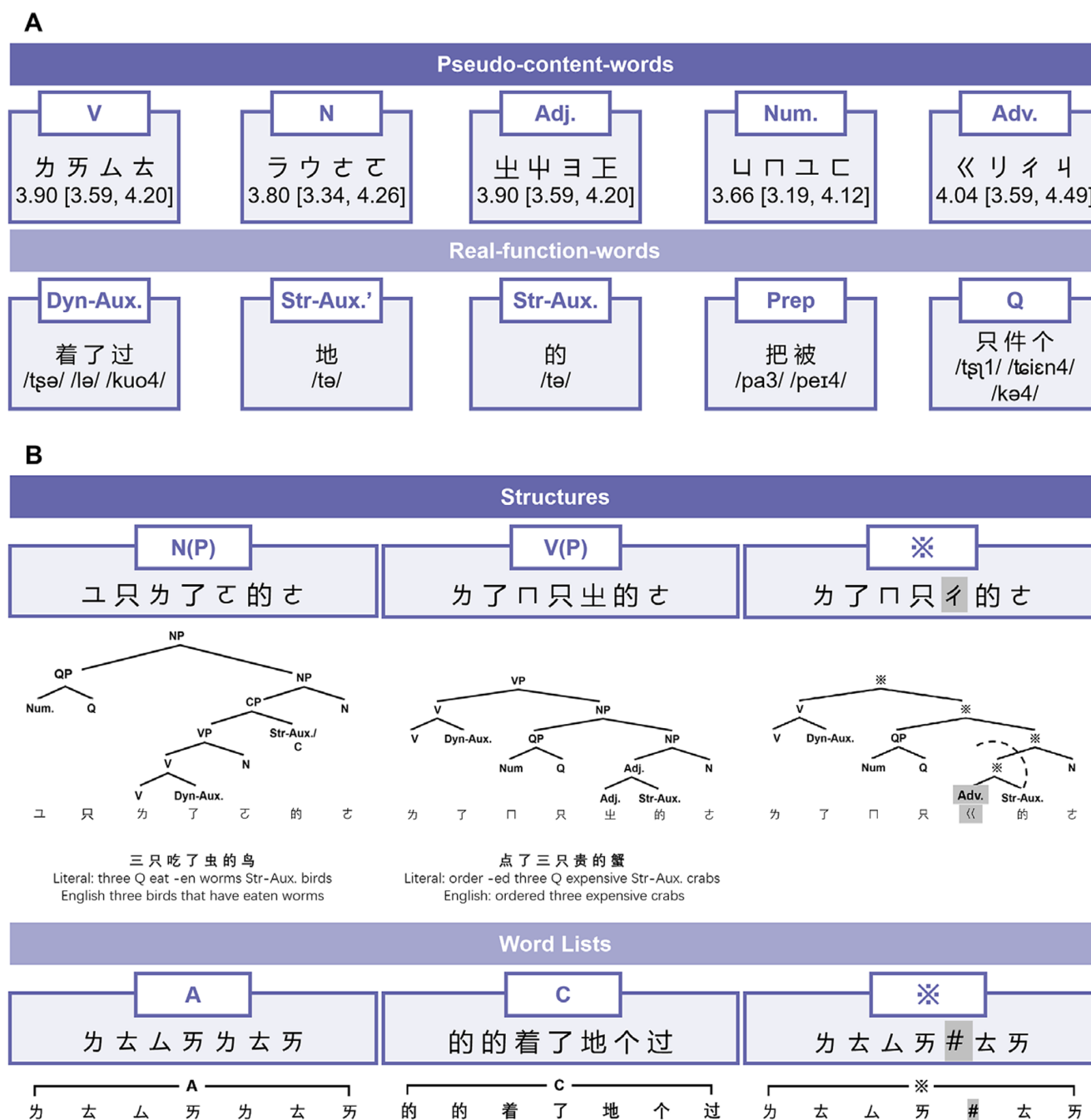
The "complex structures" were built, including 24 noun phrases and 24 verb phrases (Figure 1B), where each structure consists of three function and four pseudo-content words. See also Section 1.2 of the Supplementary Material for more details. Furthermore, in case participants could identify the phrase category by focusing only on certain function words or their combinations without processing the whole structure, 24 "Ungrammatical structures" were designed, in which the category of a content word violated the syntactic rules (see the example of "※" in Figure 1B). To avoid edge effect, the word category violations only appeared at the second or third pseudoword position. We balanced the frequency of different pseudo-content-word tokens within the same word category, the proportion of grammatical and ungrammatical trials, and the number of trials violated at different positions so as to avoid eliciting judgment preferences for participants.

To isolate the hierarchical syntactic processing effect, word lists without syntactic hierarchies were generated as a control condition (Figure 1B). The words in this condition were the same as the materials under the Structure condition to ensure that low-level lexical features would be subtracted under the contrast of "structure > word list." Using the words from the same category to avoid the possibility of syntactic merge (Zaccarella et al., 2017), we developed two types of "mono-word-category word lists": 24 lists consisting of class-specific pseudo-content word tokens, and 24 lists composed of real function words (Figure 1B). Random occurrences of "#" were set at the third to sixth positions in the rest of the 24 word list to form violations (Figure 1B), consistent with the positions where word category violations occurred in the Structure condition. The trial number of word lists and structures remained constant at 72 each, as did the length of trials, which were both 7-word sequences. Materials were pseudo-randomized and visually presented by E-prime 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA, USA).

To test the linguistic distance between Korean (L1) and Chinese (L2), we used a five-point Likert scale to develop the Chinese-Korean structure similarity scale (see Section 1.3 of the Supplementary Material). This experience/feeling-based rating approach might reflect the psychological distance at the syntactic facet and thus reflect the "structure similarity" under the psycholinguistic perspective.

### 2.3. Procedures

The experimental procedures were illustrated in Figure 2. Participants received the pseudo-content word list three days before the experiment and needed to memorize the category of each pseudoword. They were then required to accomplish a time-limited pseudoword category test by Sojump (<https://www.wjx.cn>). Only participants who scored above 90 (out of 100) were allowed to enter the formal experiment.

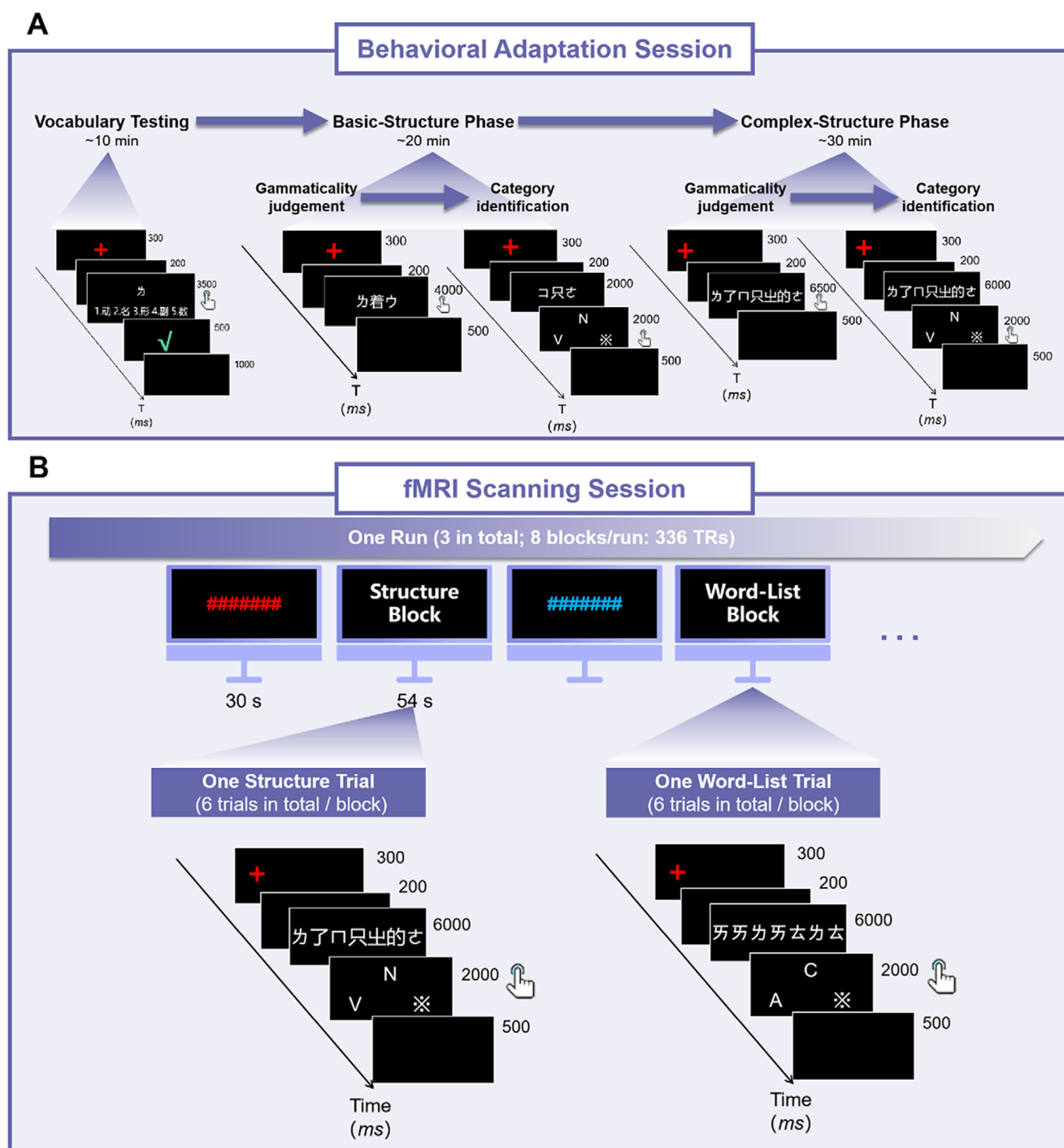


**Figure 1.** Experimental materials. (A) word categories and the corresponding tokens, including real-function-words and pseudo-content-words. Below pseudo-content-words were semantic relating mean scores with their 95% confidence intervals. (B) experimental sequences, including structures and word lists. The structure condition contained grammatical structures (noun phrases, NP and verb phrases, VP) with natural language examples provided and ungrammatical structures (marked by “※”). The word list condition contained normal lists (artificial word lists “A” and Chinese word lists “C”) and violated lists (also marked by “※”). The grey shadow marked the word category violation in the examples and the dashed curve denoted truncation of the violated part while keeping the rest of the elements still mergeable. Abbreviations: dyn-Aux.: dynamic auxiliary; str-Aux.ʹ: structural auxiliary (for verb modifiers); str-Aux.: structural auxiliary (for noun modifiers); Prep.: preposition; Q: quantifier; V: verb; N: noun; Adj.: adjective; Num.: number; Adv.: adverb. P: phrase (e.g., VP: verb phrase, NP: noun phrase); A: artificial word list; C: Chinese word list; ※: structure/word-list violation.

Before fMRI scanning, the participants underwent a behavioral adaptation phase (Figure 2A), including three sessions. Firstly, a vocabulary test, identical to the online test conducted earlier, was carried out. Participants were provided with the pseudo-content words and their corresponding syntactic categories (predefined before the experiment), and were required to map these symbols to the target syntactic categories. Only participants whose two-successive-block-accuracy reached 90% were able to pass. After that,

in the basic structure phase, they needed to complete the grammatical judgment task and the category identification task for 3-words simple structures. Finally, in the complex structure phase, participants completed the similar grammatical judgment task and the category identification task for 7-words complex structures. The presentation times and trial numbers were shown in Figure 2A.

During fMRI scanning, we employed the category identification task of complex structures under the structure condition. The



**Figure 2.** Experimental procedure. (A) Behavioral adaptation session. Participants first passed the vocabulary category identification test by the accuracy of the 2-successive-block reached 90% (20 trials per block). Then, they underwent the “basic-structure phase” to complete the grammatical judgment task (8 blocks, 8 trials per block, totaling 64 trials) and the category identification task (6 blocks, 8 trials per block, totaling 48 trials) of the basic structures. Last, in the “complex-structure phase”, participants completed the similar grammaticality judgment (16 blocks, 6 trials per block, totaling 96 trials) and category identification tasks (only two blocks of the practice section, with 6 trials each block) for the complex structure. (B) fMRI scanning session. The scanning experiment was divided into 3 runs, and 8 blocks (4 structure condition blocks and 4 word list condition blocks) were arranged in a pseudo-random manner within each run. Each block contained 6 trials and each trial was presented for 9 s, so each block lasted for 54 s. One run and the presentation of the trials with the timing parameters were shown.

experiment used a blocked-design (similar to Matchin et al., 2017) for a comparatively higher detection power, in which the structure condition and the word list condition were mixed in blocks, with different colored “#####” in the 30 s interval of the block to indicate the next task: A red one denoted the structure condition,

while a blue one indicated the word list condition. The trial numbers and the presentation time(s) were shown in Figure 2B. The whole scanning phase lasted for about 1 hour. After scanning, participants were required to complete the Chinese-Korean structure similarity test and to have a post-test interview.

## 2.4. Behavioral data analyses

Behavioral data were collected during scanning. One-sample *t*-tests were conducted to examine whether the accuracy rates of both the structure condition and word list condition significantly exceeded the chance level (i.e., 50%). As word category violations were set in the materials, if participants used task-irrelevant strategies, the accuracy rates of word category violation trials would be at random levels. Therefore, we further analyzed the accuracy rates of violated trials under the two conditions, using one-sample *t*-tests against the chance level to ensure participants' compliance with the tasks.

The effects between structure and word list (i.e., Condition) in accuracy were analyzed using a generalized linear mixed model with Accuracy as a dependent variable, Condition as a predictor, Chinese L2 proficiency scores (ChiProfS) and the linguistic distance scores measured by the Chinese-Korean structure similarity scale (LD\_SimS) as covariates to purify the main effect of Condition and Subject and Item as random variables. As for the condition effects in reaction times (RT), the linear mixed-effect model was employed that included the log-transformed RT as the dependent variable, Condition as a predictor, ChiProfS and LD\_SimS as covariates and Subject and Item as random variables. Furthermore, considering the design of the violated trials, we proceeded to employ consistent statistical analysis within each condition. Statistical models were implemented using a restricted maximum likelihood technique, following backward model-selection procedures. For instance, when constructing a generalized linear mixed model to examine the condition effect between structure and word list (SW\_condition), the initial step involves building the most complex full model with random intercepts for subjects and items, with fixed effects of independent variable (i.e., SW\_condition) and the covariate (i.e., ChiProfS and LD\_SimS). The full model was  $\text{ACC} \sim \text{SW\_Condition} + \text{LD\_SimS} + \text{ChiProfS} + (1 + \text{SW\_Condition} | \text{subject}) + (1 | \text{item})$ . As SW\_condition was a within-subject variable, it could be potentially added as a random slope and the viability was tested using *anova* function. If the full model failed to fit adequately, the model was progressively simplified by reducing random slopes until a best-fitted model was achieved. These analyses were conducted in R 4.2.2 using the package lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017).

Moreover, exploratory Pearson correlation analyses were employed to explore the relationship between accuracy, RT, ChiProfS and LD\_SimS. The multiple-comparison results were corrected via "false discovery rate" (FDR) correction.

These *t*-tests and correlation analyses were performed with SPSS 26 and jamovi 1.6.

## 2.5. Imaging data acquisition

The MR imaging data were acquired via a 3.0-Tesla Siemens PRISMA magnetic resonance scanner (Siemens AG, Erlangen, Germany) using a 64-radiofrequency-channel head coil.

For functional data acquisition, a T2\*-weighted gradient echo planar imaging (EPI) sequence was adopted with the following parameters: repetition time (TR) = 2000 ms; echo time (TE) = 30 ms; flip angle (FA) = 90°; field of view (FOV) = 208 × 208 mm<sup>2</sup>; base resolution = 104 × 104 mm<sup>2</sup>; in-plane resolution = 2 × 2 mm<sup>2</sup>; slice thickness = 2 mm; number of slices = 64; gap = 0 mm; alignment to AC-PC plane. Signals from different slices were acquired by the multi-band scanning technique (multi-band factor = 2) to efficiently minimize slice-timing effects.

Parameters for high-resolution anatomical T1-weighted images for co-registration were listed as following: TR = 2530 ms; TE = 2.27 ms; FA = 7°; FOV = 256 × 256 mm<sup>2</sup>; base resolution = 256 × 256 mm<sup>2</sup>; in-plane resolution = 1 × 1 mm<sup>2</sup>; slice thickness = 1 mm; number of slices = 208.

## 2.6. Imaging data preprocessing

The imaging data were preprocessed by DPARSF 5.1 Advanced Edition (Yan et al., 2016), implemented in the environment of MATLAB R2020b. The preprocessing steps included: (a) Removing the first 4 volumes to reduce the magnetic saturation effect; (b) Slice time correction; (c) Field mapping; (d) Spatial realignment; (e) Co-registration; (f) Segmentation (New segment + DARTEL); (g) Nuisance covariates regression (polynomial trend: 1, linear detrending), also including head motion regression by using the Friston-24 model; (h) Normalization of the images to the echo planar imaging (EPI) template based on Montreal Neurological Institute (MNI) stereotactic space to minimize cerebral differences between participants, and resampled the images into 2 × 2 × 2 mm<sup>3</sup>; (i) Smoothing the images with a 3D Gaussian kernel with full-width at half-maximum (FWHM) of 4 mm.

## 2.7. Whole-brain level analyses

The whole-brain level analyses were conducted using SPM 12 (<https://www.fil.ion.ucl.ac.uk/spm/>). At the first level, a general linear model (GLM) was established for each participant by including the structure and the word list conditions as two regressors of interest, with the onset and duration (54 s) of each block modulated as a boxcar function convolved with a canonical hemodynamic response function. The data were further high-pass filtered at 128 Hz to eliminate low-frequency drifts. The "structure > word list" contrast results of each participant were then entered into the second-level analysis.

At the second-level group analysis, we performed a one-sample *t*-test to examine whether there was a significant difference between the structure and word list conditions. Each individual's mean FD (framewise displacement) Jenkinson value accounted for head motion artifacts, accuracy, reaction times, ChiProfS and LD\_SimS were included as covariates and regressed out. The results of the whole-brain analysis were reported using a FWE (family-wise error)-corrected threshold of  $p < .05$ , cluster size ( $K_E$ ) ≥ 20. Furthermore, we detected activation under each condition separately ( $p_{\text{uncorrected}} < .00001$ ,  $K_E \geq 50$ ) in contrast to the implicit baseline (i.e., 0) to ensure normal processing in both conditions following Chen et al. (2023b).

## 2.8. Region of interest (ROI) analyses

Consistent with the previous study (Chen et al., 2023b), we utilized a functional left-hemispheric language atlas based on 220 participants, developed by Fedorenko et al. (2010) (<http://web.mit.edu/evlab/funclab/>), to investigate the activation differences under the "structure > word list" contrast within the language network. This atlas was employed as the language mask for small volume correction (SVC) to identify peak activity coordinates (cluster-level  $p_{\text{FWE}} < .05$ , using the cluster defining threshold at  $p_{\text{uncorrected}} < .001$ ,  $K_E \geq 30$ ).

The percentage of signal change was measured within each region of interest (ROI) based on the original BOLD (Blood Oxygenation Level Dependent) signal, without regressing out the

behavioral indices (i.e., accuracy, RT, ChiProfS and LD\_SimS). The correlations between the percentage of signal change and the behavioral indices were further examined.

## 2.9. Effective connectivity analysis

To further investigate the information transfer among the ROIs, we employed the unified structural equation modeling (uSEM) approach for the effective connectivity analysis (Gates et al., 2011) following Chen et al. (2023b). The original signal from each ROI under the structure condition were extracted, without regressing out the behavioral indices. These signals were then included in Group Iterative Multiple Model Estimation (a MATLAB-compatible toolkit, GIMME, Gates et al., 2011) for model specification. GIMME estimates the optimal models from the group-level to the individual-level using Lagrange multiplier tests. If a connection significantly improved the model fit ( $\geq 75\%$  of the sample), it was added to the model for re-estimation. This uSEM approach can provide more detailed time information, including lagged effects (i.e., effects from a previous time point to the current one), and contemporaneous effects (i.e., effects at the same time point) (Gates et al., 2011).

Subsequently, we extracted the effective connectivity strength of each connection between the ROIs and then performed the exploratory Pearson correlation tests between each connectivity strength and the behavioral indices to evaluate the modulative impact of learner factors.

## 3. Results

### 3.1. Behavioral results

The descriptive behavioral results are shown in Table 1. The accuracy rate of the structure condition was significantly higher than the chance level ( $t(19) = 4.52, p < .001, d = 1.01$ ) and did not differ statistically from 65% ( $t(19) = -.68, p = .51$ ), indicating that the participants were compliant with the experiment task (with standard of 65%, referring to Iwabuchi & Makuuchi, 2021). Additionally, the accuracy rate of ungrammatical trials also surpassed the chance level ( $t(19) = 4.12, p < .001, d = .92$ ), suggesting that participants processed without relying on task-unrelated strategies. The word list condition showed a similar pattern, with accuracy rate significantly above chance levels ( $ts(19) \geq 21.5, ps < .001, ds \geq 4.81$ ).

The results for accuracy and reaction times analyses using statistical models were shown in Tables 2 and 3. To note, as the full models fitted best, all the subsequent analyses were conducted using full models, as detailed in the footnotes in Tables 2 and 3. The structure condition had lower accuracy than the word list condition (Estimate =  $-2.68, SE = .30, z = -8.94, p < .001$ ), along with longer RT (Estimate =  $.08, SE = .01, t = 5.77, p < .001$ ). These findings indicated that the structure condition posed a greater processing

challenge than the word list condition. Therefore, both accuracy and RT were included as covariance in defining the group-level design matrix.

Further analysis under structure condition revealed no statistical difference in accuracy between the grammatical and the ungrammatical structures (Estimate =  $.22, SE = .21, z = 1.06, p = .291$ ), suggesting that both types of structures were processed effectively without resorting to task-irrelevant strategies. Moreover, participants took significantly longer to process grammatical structures (Estimate =  $-.06, SE = .01, t = -5.47, p < .001$ ). This could be due to the fact that grammatical structures need to be processed to the very end so as to merge the whole structures, whereas ungrammatical structures can be detected immediately upon word category violation.

The analysis under the word list condition also yielded similar results. No accuracy differences between the normal and violated word lists were reported (Estimate =  $-.02, SE = .30, z = -.08, p = .935$ ). However, there were no significant differences in RT between the normal and violated word lists either (Estimate =  $-.01, SE = .01, t = -1.12, p = .277$ ). This may be due to distinct processing strategies: in the normal word list condition, hierarchical processing was not required, resulting in similar detection times for normal and violated trials.

As for the L2 learner factors, the Pearson correlation analyses showed that ChiProfS significantly correlated with the accuracy of the structure condition ( $r = .52, p = .02$ ). Further linear regression analysis revealed that ChiProfS accounted for 26.8% of its variation ( $R^2 = .268$ ). The regression model was significant ( $F(1, 18) = 6.60, p = .02$ ), with ChiProfS emerging as a significant predictor ( $\beta = .52, t = 2.57, p = .02$ ). However, the accuracy under the word list condition was not significantly correlated with ChiProfS ( $r = .36, p = .12$ ). And LD\_SimS ( $M = 2.86, SD = 0.77$ ; CI:  $[2.49, 3.23]$ ) were not significantly correlated with the accuracy in either condition ( $rs \geq -.25, ps > .05$ ).

### 3.2. Whole-brain level results

Each condition showed reliable activation when compared with the implicit baseline, guaranteeing that both structure and word list conditions were normally processed (Table 4 and Figure 3A). Under the single structure condition, a large cluster containing the left MFG and IFG was detected. While under the contrast “structure > word list” condition, the significant activity of the superior parietal lobule (SPL) was found (Table 4 and Figure 3B<sub>1</sub>).

### 3.3. ROI analyses results

By employing the functional language atlas under the “structure > word list” contrast, the SVC results identified two peak-activity coordinates located in the left pIFG ( $-56, 26, 16$ ) and the left precentral gyrus (PreCG;  $-48, 4, 46$ ) (see Table 4 and Figure 3B<sub>2</sub>). A

**Table 1.** Behavioral results

| Conditions          | Structure        |                    |                  |                  | Word list       |                  |                 |                 |
|---------------------|------------------|--------------------|------------------|------------------|-----------------|------------------|-----------------|-----------------|
|                     | Mean (SD)        | 95% CI             | Gram. (SD)       | Ungram. (SD)     | Mean (SD)       | 95% CI           | Normal (SD)     | Violated (SD)   |
| Accuracy            | 0.63 (0.13)      | [0.57, 0.69]       | 0.62 (0.13)      | 0.66 (0.17)      | 0.92 (0.07)     | [0.89, 0.95]     | 0.91 (0.08)     | 0.93 (0.07)     |
| Reaction times (ms) | 1095.84 (119.74) | [1038.35, 1153.34] | 1162.40 (143.99) | 1001.05 (139.01) | 906.84 (146.01) | [836.73, 976.95] | 919.55 (158.24) | 884.85 (141.43) |

Abbreviations: Gram.: Grammatical structure; Ungram.: Ungrammatical structure; SD: standard deviation; CI: confidence intervals.

**Table 2.** The generalized linear mixed model results for accuracy

| Parameters                         | Predictor             | Fixed effects |      |         |           | Random effects |      |          |         |
|------------------------------------|-----------------------|---------------|------|---------|-----------|----------------|------|----------|---------|
|                                    |                       | Estimate      | SE   | z value | p         | Subject        |      | Item     |         |
|                                    |                       |               |      |         |           | Variance       | SD   | Variance | SD      |
| Accuracy in total                  | (Intercept)           | 2.98          | 1.42 | 2.11    | 0.035*    | 0.22           | 0.47 | 0.37     | 0.61    |
|                                    | Condition S-W         | −2.68         | 0.30 | −8.94   | < .001*** |                |      |          |         |
|                                    | LD_SimS               | −0.33         | 0.17 | −1.89   | 0.059     |                |      |          |         |
|                                    | ChiProfS              | 0.07          | 0.04 | 1.57    | 0.117     |                |      |          |         |
| Accuracy under structure condition | (Intercept)           | −1.34         | 1.57 | −0.85   | 0.394     | 0.27           | 0.52 | 0.26     | 0.51    |
|                                    | Violation Gram-Ungram | 0.22          | 0.21 | 1.06    | 0.291     |                |      |          |         |
|                                    | LD_SimS               | −0.11         | 0.19 | −0.61   | 0.540     |                |      |          |         |
|                                    | ChiProfS              | 0.08          | 0.05 | 1.70    | 0.090     |                |      |          |         |
| Accuracy under word list condition | (Intercept)           | 3.16          | 2.54 | 1.25    | 0.213     | 1.04           | 1.02 | 5.15E−8  | 2.27E−4 |
|                                    | Violation Norm-Viol   | −0.02         | 0.30 | −0.08   | 0.935     |                |      |          |         |
|                                    | LD_SimS               | −0.33         | 0.32 | −1.04   | 0.297     |                |      |          |         |
|                                    | ChiProfS              | 0.02          | 0.08 | 0.23    | 0.817     |                |      |          |         |

The generalized linear mixed models were as follows:

ACC~SW\_Condition + LD\_SimS + ChiProfS + (1 + SW\_Condition | subject) + (1 | item);

ACC~ Gram-Ungram\_Condition + LD\_SimS + ChiProfS + (1 + Gram-Ungram\_Condition | subject) + (1 | item);

ACC~ Norm-Viol\_Condition + LD\_SimS + ChiProfS + (1 + Norm-Viol\_Condition | subject) + (1 | item).

Abbreviations: Condition S-W: The condition effects of comparing structure conditions and word list conditions; ChiProfS: Chinese L2 proficiency scores; LD\_SimS: linguistic distance indexed by the similarity scores of the structures; Violation Gram-Ungram: The violation effects of comparing grammatical and ungrammatical structures; Violation Norm-Viol: The violation effects of comparing normal and violated word lists. Significance levels: \* =  $p < .05$ ; \*\* =  $p < .01$ ; \*\*\* =  $p < .001$ .

**Table 3.** The linear mixed-effect model results for reaction times (logRT)

| Parameters                   | Predictor             | Fixed effects |         |         |           | Random effects |      |          |      |
|------------------------------|-----------------------|---------------|---------|---------|-----------|----------------|------|----------|------|
|                              |                       | Estimate      | SE      | t value | p         | Subject        |      | Item     |      |
|                              |                       |               |         |         |           | Variance       | SD   | Variance | SD   |
| RT in total                  | (Intercept)           | 2.88          | 0.14    | 20.96   | < .001*** | 5.24E−3        | 0.07 | 7.16E−4  | 0.03 |
|                              | Condition S-W         | 0.08          | 0.01    | 5.77    | < .001*** |                |      |          |      |
|                              | LD_SimS               | 0.01          | 0.02    | 0.78    | 0.448     |                |      |          |      |
|                              | ChiProfS              | 7.54E−4       | 4.20E−3 | 0.18    | 0.859     |                |      |          |      |
| RT under structure condition | (Intercept)           | 2.96          | 0.13    | 22.02   | < .001*** | 2.11E−3        | 0.05 | 5.28E−4  | 0.02 |
|                              | Violation Gram-Ungram | −0.06         | 0.01    | −5.47   | < .001*** |                |      |          |      |
|                              | LD_SimS               | 0.02          | 0.02    | 0.98    | 0.340     |                |      |          |      |
|                              | ChiProfS              | 1.23E−3       | 4.12E−3 | 0.30    | 0.769     |                |      |          |      |
| RT under word list condition | (Intercept)           | 2.87          | 0.20    | 14.59   | < .001*** | 6.00E−3        | 0.08 | 1.85E−4  | 0.01 |
|                              | Violation Norm-Viol   | −0.01         | 0.01    | −1.12   | 0.277     |                |      |          |      |
|                              | LD_SimS               | −2.13E−4      | 0.02    | −0.01   | 0.993     |                |      |          |      |
|                              | ChiProfS              | 2.46E−3       | 6.04E−3 | 0.41    | 0.689     |                |      |          |      |

The linear mixed-effect models were as follows:

RT ~ SW\_Condition + LD\_SimS + ChiProfS + (1 + SW\_Condition | subject) + (1 | item);

RT ~ Gram-Ungram\_Condition + LD\_SimS + ChiProfS + (1 + Gram-Ungram\_Condition | subject) + (1 | item);

RT ~ Norm-Viol\_Condition + LD\_SimS + ChiProfS + (1 + Norm-Viol\_Condition | subject) + (1 | item).

Abbreviations: Condition S-W: The condition effects of comparing structure conditions and word list conditions; ChiProfS: Chinese L2 proficiency scores; LD\_SimS: linguistic distance indexed by the similarity scores of the structures; Violation Gram-Ungram: The violation effects of comparing grammatical and ungrammatical structures; Violation Norm-Viol: The violation effects of comparing normal and violated word lists. Significance levels: \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ .

semantic deactivation pattern was also detected within the language atlas, see Section 2 of the Supplementary Material. Since the significantly activated SPL is not included in the functional language atlas, we incorporated SPL into our ROI analyses.

We extracted the percentage of signal change for three ROIs (pIFG, PreCG and SPL) and conducted separate Pearson correlation tests between the signal intensity and behavioral indices to explore their neuro-cognitive relationships for each ROI (thus with

**Table 4.** Whole-brain level and ROI-level results

| Contrasts                                                               | Region               | K <sub>E</sub> | MNI peak coordinates (mm) |     |     | T-value |
|-------------------------------------------------------------------------|----------------------|----------------|---------------------------|-----|-----|---------|
|                                                                         |                      |                | x                         | y   | z   |         |
| Whole-brain level ( $p_{\text{uncorrected}} < .00001$ , $K_E \geq 50$ ) |                      |                |                           |     |     |         |
| Structure > 0                                                           | RCerebellum exterior | 3147           | 8                         | −80 | −26 | 16.75   |
|                                                                         |                      |                | 30                        | −58 | −34 | 14.48   |
|                                                                         |                      |                | −28                       | −90 | 8   | 14.32   |
|                                                                         | LSPL                 | 744            | −28                       | −70 | 30  | 12.98   |
|                                                                         |                      |                | −38                       | −44 | 32  | 12.32   |
|                                                                         |                      |                | −34                       | −48 | 38  | 11.51   |
|                                                                         | LThalamus proper     | 128            | −16                       | −22 | 6   | 12.94   |
|                                                                         |                      |                | −20                       | −18 | 18  | 8.09    |
|                                                                         | LMFG/IFG             | 422            | −26                       | 2   | 52  | 12.87   |
|                                                                         |                      |                | −48                       | 4   | 44  | 11.55   |
|                                                                         |                      |                | −42                       | 2   | 50  | 11.13   |
|                                                                         | LSMC                 | 333            | −6                        | 12  | 46  | 12.40   |
|                                                                         |                      |                | 8                         | 10  | 42  | 9.04    |
|                                                                         |                      |                | 0                         | 24  | 40  | 8.28    |
|                                                                         | RPHG                 | 61             | 20                        | −34 | −10 | 12.24   |
|                                                                         |                      |                | 20                        | −26 | −12 | 11.19   |
|                                                                         |                      |                | 26                        | −32 | −2  | 10.67   |
|                                                                         | RSOG                 | 231            | 30                        | −78 | 40  | 11.25   |
|                                                                         |                      |                | 32                        | −62 | 34  | 9.65    |
|                                                                         |                      |                | 28                        | −62 | 48  | 8.27    |
|                                                                         | LCaudate             | 87             | −16                       | −4  | 14  | 10.23   |
|                                                                         |                      |                | −14                       | 6   | 4   | 8.47    |
|                                                                         |                      |                | −22                       | −8  | 8   | 8.08    |
|                                                                         |                      | 190            | −30                       | −70 | −32 | 10.09   |
|                                                                         |                      |                | −26                       | −64 | −38 | 9.96    |
|                                                                         |                      |                | −44                       | −68 | −32 | 9.33    |
|                                                                         |                      | 63             | 16                        | −18 | 3   | 9.47    |
|                                                                         |                      |                | 14                        | −26 | 6   | 8.67    |
|                                                                         |                      |                | 75                        | −16 | −96 | 0       |
|                                                                         |                      | 84             | −22                       | −30 | −8  | 9.01    |
|                                                                         |                      |                | −26                       | −22 | −10 | 7.97    |
|                                                                         |                      |                | −30                       | −32 | −6  | 6.52    |
| Word list > 0                                                           | LCalc                | 1343           | −14                       | −78 | 4   | 11.88   |
|                                                                         |                      |                | 0                         | −90 | −8  | 11.36   |
|                                                                         |                      |                | −36                       | −76 | −18 | 11.12   |
|                                                                         | LSPL                 | 245            | −26                       | −60 | 44  | 10.64   |
|                                                                         |                      |                | −24                       | −60 | 34  | 10.39   |
|                                                                         |                      |                | −34                       | −44 | 34  | 9.80    |
|                                                                         | LSMC                 | 61             | −8                        | 8   | 48  | 10.34   |
|                                                                         |                      |                | −4                        | 14  | 42  | 7.18    |
|                                                                         | RIOG                 | 81             | 36                        | −84 | −10 | 8.61    |

(Continued)

Table 4. (Continued)

| Contrasts                                                                 | Region | $K_E$ | MNI peak coordinates (mm) |     |     | T-value |
|---------------------------------------------------------------------------|--------|-------|---------------------------|-----|-----|---------|
|                                                                           |        |       | x                         | y   | z   |         |
|                                                                           |        |       | 38                        | -70 | -20 | 8.39    |
|                                                                           |        |       | 32                        | -72 | -14 | 8.16    |
|                                                                           | LIOG   | 62    | -30                       | -88 | -2  | 7.80    |
|                                                                           |        |       | -22                       | -92 | -2  | 6.94    |
|                                                                           |        |       | -28                       | -96 | 4   | 6.78    |
| Whole-brain level ( $p_{FWE} < .05$ , $K_E \geq 20$ )                     |        |       |                           |     |     |         |
| Structure > word list                                                     | SPL    | 22    | -26                       | -72 | 40  | 11.95   |
| ROI-level (within the language atlas) ( $p_{FWE} < .05$ , $K_E \geq 30$ ) |        |       |                           |     |     |         |
| Structure > word list                                                     | LPreCG | 218   | -48                       | 4   | 46  | 10.25   |
|                                                                           | LpIFG  | 473   | -56                       | 26  | 16  | 7.12    |

Abbreviations: RCerebellum exterior: right cerebellum exterior; LSPL: left superior parietal lobule; LThalamus Proper: left thalamus proper; LMFG: left middle frontal gyrus; IFG: inferior frontal gyrus; LSMC: left supplementary motor cortex; RPHG: right parahippocampal gyrus; RSOG: right superior occipital gyrus; LCaudate: left caudate; LCalc: left calcarine cortex; RIOG: right inferior occipital gyrus; LIOG: left inferior occipital gyrus; LPreCG: left precentral gyrus; LpIFG: left posterior inferior frontal gyrus.  $K_E$ : cluster size.

$p$ -value uncorrected). The results, as shown in Table 5 and Figure 4, revealed significant positive correlations between the signal change percentages of all three ROIs and accuracy under the structure condition ( $r_s \geq .45$ ,  $p_{\text{uncorrected}} \leq .05$ ). Furthermore, the signal intensity of SPL showed a significant positive correlation with ChiProfS ( $r = .59$ ,  $p = .01$ ), whereas the signal intensity of PreCG exhibited a significant negative correlation with reaction times under the structure condition ( $r = -.49$ ,  $p = .03$ ).

### 3.4. Effective connectivity modeling results

Model fit indices indicated that the group-level effective connectivity model was reliably estimated (CFI = 1.00, NNFI = 1.00) (see also Chen et al., 2023b; Gates et al., 2011). The uSEM results revealed a neural circuit, in which (a) the left pIFG projected a contemporaneous connection to the left PreCG; (b) the PreCG further projected contemporaneous connections to the left SPL and (c) the SPL transferred the information back to pIFG via a contemporaneous connection. All these results were shown in Figure 3C.

Regarding Pearson correlation tests between each connectivity strength and the behavioral indices, the results indicated that (Table 5 and Figure 4) the strength of the contemporaneous PreCG-to-SPL connection exhibited a significant negative correlation with LD\_SimS ( $r = -.46$ ,  $p_{\text{uncorrected}} < .05$ ).

## 4. Discussion

The present fMRI study aimed to investigate the neural mechanisms underlying Chinese L2 syntactic processing. Significant activation in the left SPL at the whole-brain level was revealed, as well as in the left pIFG and PreCG within the language network. Based on these ROIs, we localized a frontoparietal network that might play a crucial role in Chinese L2 syntactic processing. Regarding the L2 learner factors, we found a significant correlation between ChiProfS and both behavioral accuracy and the activation intensity of the SPL, respectively. Moreover, LD\_SimS showed a significant correlation with the strength of contemporaneous connectivity from PreCG to the SPL. Thus, these findings indicated that Chinese L2 syntactic processing shared the core neural basis in the frontal

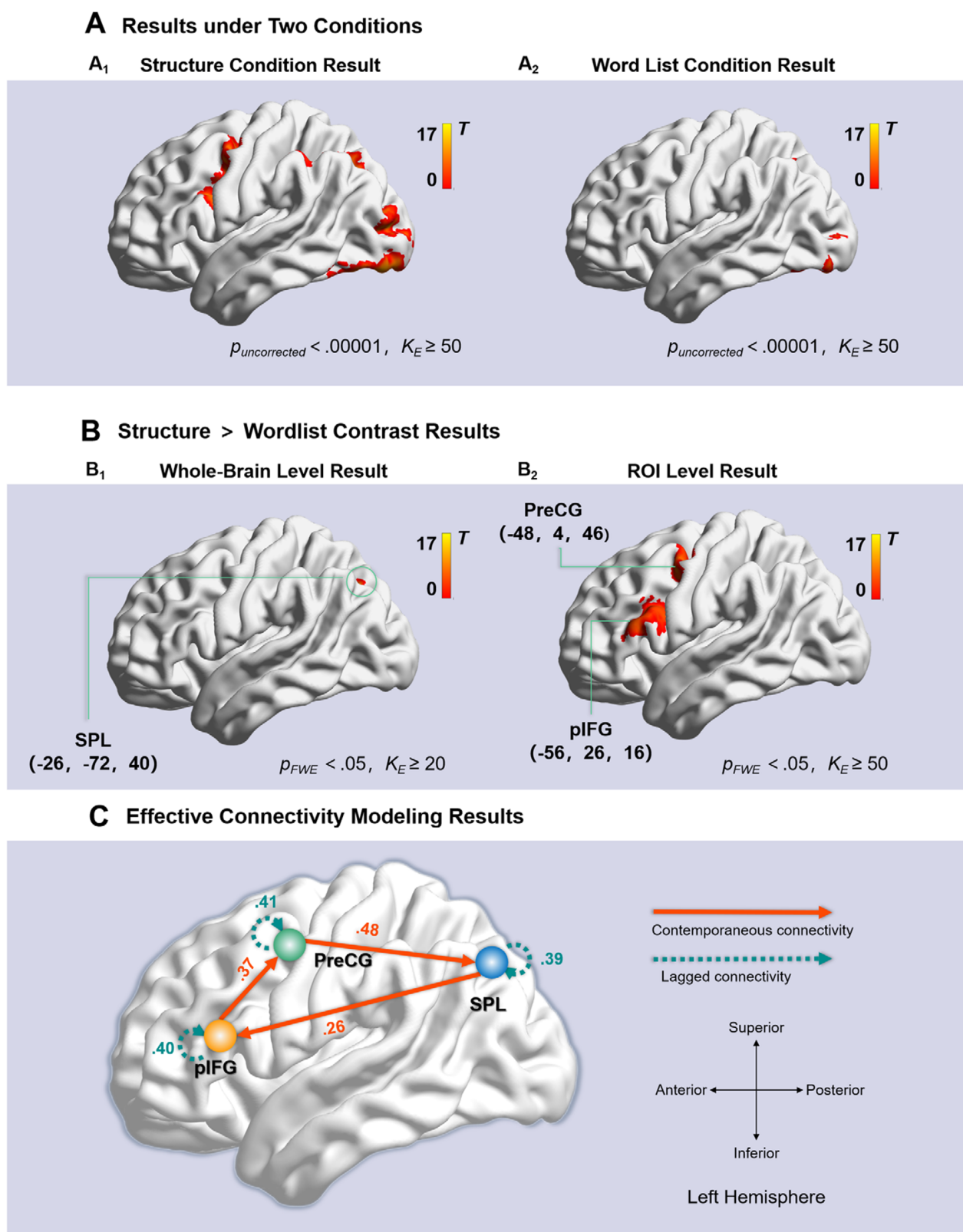
cortex (pIFG and PreCG) regardless of language typological differences and learner characteristics and, more importantly, recruited a crucial frontoparietal network which was modulated by L2 learner factors. Therefore, the present study provided further neurobiological evidence for certain second language processing hypothesis, such as the SSH.

### 4.1. SPL: the critical region for L2 syntactic processing

Significant involvement of the SPL in L2 learning was previously reported (Lee et al., 2007; Mechelli et al., 2004; Xiang et al., 2012). For instance, Mechelli et al. (2004) found that bilinguals had significantly higher gray matter density in the left parietal cortex, including the SPL, compared to monolinguals, and that gray matter density was correlated with L2 proficiency positively. Consistently, the intensity of SPL was positively correlated with both ChiProfS and accuracy in complex syntactic processing in this study, reflecting the intimate relationship between SPL and L2 learning.

Gray matter density in the SPL was strongly and positively associated with lexical learning and semantic integration (Lee et al., 2007). Xiang et al. (2012) revealed that the pathway connecting SPL and frontal lobes was linked to rapid vocabulary acquisition. The SPL was assumed to support explicit learning strategies that facilitated the association of novel words with concepts. However, in this study, semantic information was diminished, and the significant activation was found only in the SPL of L2 learners compared to L1 speakers (Chen et al., 2023b). This indicated that the SPL also contributed to L2 syntactic processing, expanding our understanding of its functional involvement.

It is worth noting that Xiang et al. (2012) discussed the relationship between the SPL and explicit learning strategies. Yang and Li (2012) utilized an artificial grammar learning paradigm to investigate the neural differences between explicit and implicit learning. The learning process was completed before fMRI scanning. The results revealed that only the explicit learning group significantly activated the SPL. This is in line with the notion that the SPL might be involved in the process of explicit representation of syntactic rules (Seger et al., 2000). Therefore, it seemed that Chinese L2 learners had utilized explicit Chinese syntactic rules to complete the syntactic tasks, thereby activating the SPL in this study.

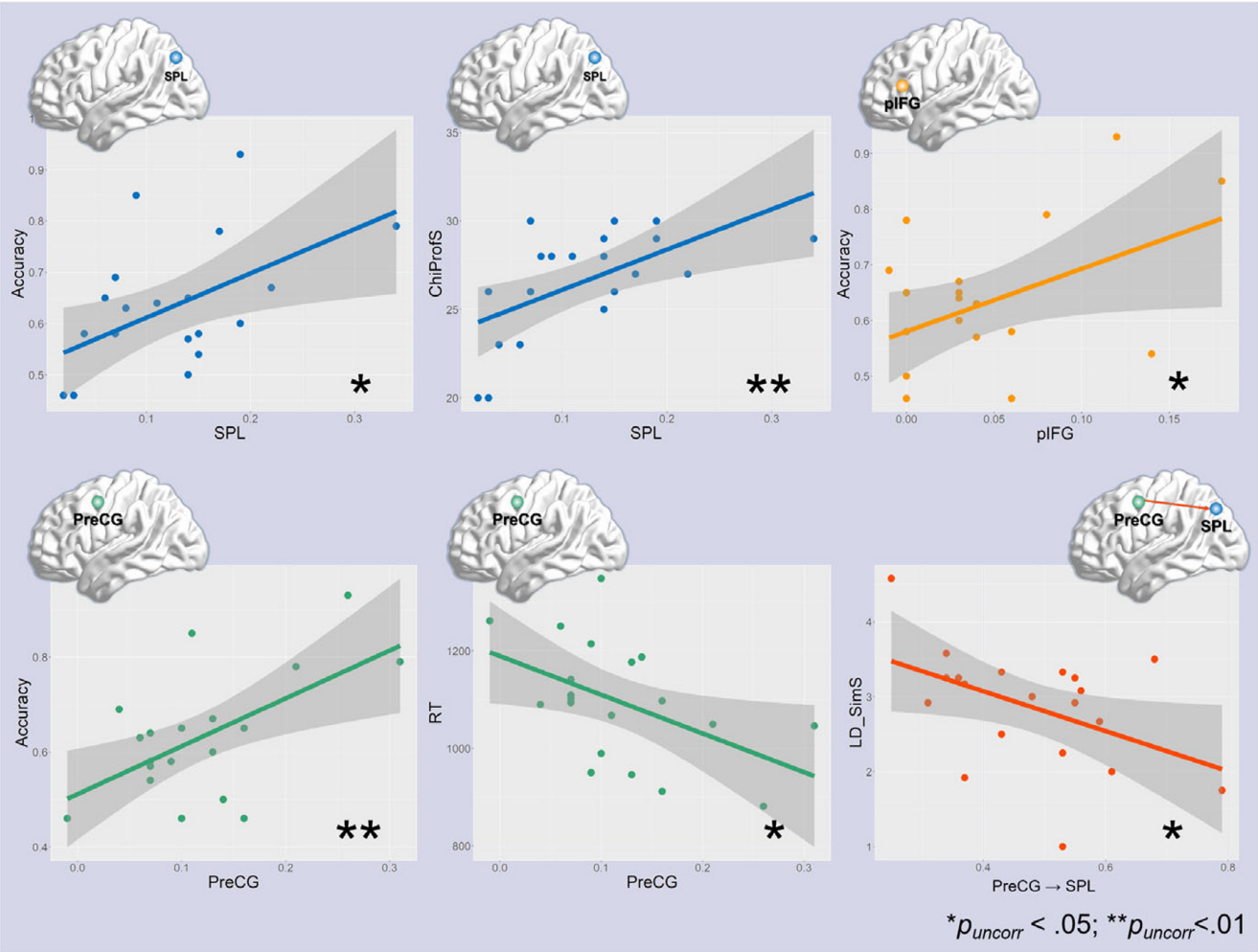


**Figure 3.** Imaging results. (A) The whole-brain level result for each condition. (B) “structure > word list” results: B<sub>1</sub>: at the whole-brain level; B<sub>2</sub>: the small volume correction result at the ROI level (2 regions of interest were identified). (C) Effective connectivity modeling results via uSEM. Group-mean connectivity strength (i.e., beta value) was also presented for each connection. Abbreviations: SPL: superior parietal lobule; PreCG: precentral gyrus; pIFG: posterior inferior frontal gyrus.  $K_E$ : cluster size.

**Table 5.** Correlation tests between the signal intensity, the connectivity strength and the behavioral indices

| Correlations                          |               | Behavioral indices |        |               |         |
|---------------------------------------|---------------|--------------------|--------|---------------|---------|
|                                       |               | Accuracy           | RT     | ChiProfS      | LD_SimS |
| Signal change%                        | pIFG          | <b>.452*</b>       | −.224  | .292          | −.293   |
|                                       | PreCG         | <b>.587**</b>      | −.486* | .251          | .060    |
|                                       | SPL           | <b>.525*</b>       | −.037  | <b>.589**</b> | −.066   |
| Contemporaneous connectivity strength | pIFG to PreCG | .114               | −.131  | −.215         | .256    |
|                                       | PreCG to SPL  | .393               | −.350  | .353          | −.462*  |
|                                       | SPLto pIFG    | .042               | .395   | .279          | −.019   |
| Lagged connectivity strength          | pIFG to pIFG  | .080               | −.144  | −.183         | .208    |
|                                       | PreCGto PreCG | −.191              | .123   | −.170         | .185    |
|                                       | SPL to SPL    | .247               | −.001  | .153          | .240    |

Abbreviations: pIFG: the left posterior inferior frontal gyrus; PreCG: precentral gyrus; SPL: the left superior parietal lobule; accuracy: the accuracy rate under the structure condition; RT: reaction times under the structure condition; ChiProfS: Chinese L2 proficiency scores; LD\_SimS: linguistic distance indexed by the similarity scores of the structures; \* $p_{uncorrected} < .05$ , \*\* $p_{uncorrected} < .01$ ; Statistically significant values were bolded.



**Figure 4.** Correlation results. Correlation results between the signal intensity (or the connectivity strength) and the behavioral indices. Abbreviations: SPL: the left superior parietal lobule; pIFG: the left posterior inferior frontal gyrus; PreCG: precentral gyrus; Accuracy: accuracy rate under the structure condition; RT: reaction times under the structure condition; ChiProfS: Chinese L2 proficiency scores; LD\_SimS: linguistic distance indexed by the similarity scores of the structures; PreCG → SPL: contemporaneous connectivity strength of PreCG-to-SPL; \* $p_{uncorrected} < .05$ , \*\* $p_{uncorrected} < .01$ .

Nevertheless, backing up L2 syntactic processing, by no means, inferred that SPL is an L2 syntax-specific region, but rather SPL has been proposed to be a key region within the multiple demand (MD) network (Blank & Fedorenko, 2017; Chen et al., 2021; Diachek et al., 2020; Fedorenko, 2014; MacGregor et al., 2022). A large-scale fMRI investigation showed that in contrast with the language-selective frontotemporal network, the domain-general MD network, including bilateral frontoparietal areas, exhibited a stronger response in experiments with explicit tasks than in passive reading/listening paradigms (Diachek et al., 2020). This suggested that the engagement of the MD network during language processing reflected cognitive effort associated with extraneous task demands. SPL might house the general cognitive abilities, such as attention, working memory and spatial processing including spatial attention.

On the one hand, the use of explicit rules largely relied on consciousness and attention control. Previous research had indicated that the activation in the left superior parietal region (including the SPL) was associated with enhanced attention during high-level language monitoring (Geva, 2023). Therefore, this study found significant activation in the SPL indicating that L2 learners might allocate more attention resources in performing complex syntactic processing tasks. On the other hand, L2 syntactic processing might rely more on basic cognitive processes such as working memory, resulting in greater working memory demands (Hou et al., 2024; McDonald, 2006). Friederici (2017: 47–49) proposed that complex language processing might also involve the non-syntactic-specific verbal working memory from the parietal lobe. Moreover, Koenigs et al. (2009) found that patients with the SPL damage exhibited deficits in manipulating and rearranging information in working memory. Nečka et al. (2021) further reported that working memory training enhanced the correlation between task performance accuracy and activation of the SPL subserving working memory. These findings suggested that the SPL might play a crucial role in working memory.

Furthermore, in addition to verbal working memory, research has highlighted the important role of the SPL in spatial processing, including spatial attention (Greenberg et al., 2010; Ritz & Shenhav, 2024; Szczepanski et al., 2010). For instance, Greenberg et al. (2010), using a visual attention task involving the shifting of stimulus position and color, found that both spatial shifts of attention and (nonspatial) color shifts elicited significant activation in the medial SPL. This indicated that the medial SPL served as a cortical hub for initiating attention and task shifts across multiple domains. Similarly, in studies on letter string and word processing during reading, researchers also reported significant SPL activation (Lobier et al., 2012; Seger et al., 2000; Sun et al., 2011; Wu et al., 2012), suggesting that attentional mechanisms supported by the SPL were involved in visual word recognition. For instance, Seger et al. (2000) observed enhanced SPL activation when the surface structure of the strings changed in an artificial grammar judgment task. This shift involved forming a mapping between the original letter set and the transferred letter set, reflecting not only the explicit syntactic rule processing discussed earlier but also the SPL's role in spatial reasoning. In the present study, given that both the structure and word-list conditions contained the same visual symbols, the contrast of “structure > word list” was expected to mitigate the effects from lower-level processing of spatial features. However, as a key node in the multi-demand network, SPL detected in the present study might also be multi-functional; that is, it is hard to completely exclude the differences in the amount of spatial attention and executive control (Osaka et al., 2004; Shomstein, 2012; Sulpizio

et al., 2023) allocated to different sequence conditions. Nevertheless, the activation of SPL demonstrated that L2 syntactic processing might recruit more general cognitive brain regions beyond the scope of the language network (Hou et al., 2024). The SPL activation identified in this study could be associated with higher memory load, attention demands, or other executive control cognitive abilities in L2 learners for coping with the challenging syntactic processing task. Future studies may design a comparable spatial processing task to compare with the language (esp., syntactic) tasks to specify the functional role of SPL.

#### 4.2. The crucial frontoparietal network for L2 syntactic processing

The ROI analyses of the present study indicated that the left posterior inferior frontal gyrus (pIFG) and the left precentral gyrus (PreCG) together with the left SPL formed a functional network. We further performed the uSEM to analyze the effective connectivity between these regions, which pointed toward the existence of a functional neural circuit that may back up the L2 syntactic processes.

The left posterior inferior frontal gyrus (pIFG), which roughly corresponded to the Broca's area, served as a core brain region for syntactic processing. As Zaccarella et al. (2017) pointed that, Brodmann Area 44 in the pIFG functioned as a Merge engine to support hierarchical syntactic processing. It also played a crucial role in second language learning and processing (Friederici, 2017; Sakai et al., 2004). Additionally, the study utilizing transcranial direct current stimulation (tDCS) had demonstrated the causal role of this region with the activation of this area being causally related to L2 learning outcomes (de Vries et al., 2010). This study found a positive correlation between the activation intensity of pIFG and the accuracy under complex syntactic processing, confirming the significant role of pIFG in Chinese L2 syntactic processing and indicating shared neural foundations with Chinese L1 syntactic processing (Chen et al., 2023b).

PreCG, as a component of the “phonological loop” with the IFG, played a crucial role in inner speech rehearsal (Kaestner et al., 2022; Wheat et al., 2010). Wheat et al. (2010), using MEG in a pseudo-homophone priming task, observed stronger responses to pseudo-homophone priming compared to orthographic priming of visually presented words within 100 ms of target word onset. These responses were localized to a cluster that included the left IFG and PreCG, confirming the role of the PreCG in the early phonological processing of visual word recognition. Furthermore, Kaestner et al. (2022), employing cortical electrophysiology, highlighted the role of the PreCG in transducing visual orthographic information into auditory phonological codes during silent reading. Here, the present study revealed a positive correlation between the activation intensity of PreCG and accuracy, and a negative correlation with RT, suggesting that the stronger the activation of PreCG, the better the behavioral performance of L2. Bernal and Ardila (2009) pointed out the connection between the PreCG and the Broca's area (BA 44) via the arcuate fasciculus (AF), which played an important role in speech encoding, providing a structural basis for language imitation and serving as the first step in language learning and acquisition. In our behavioral adaptation stages prior to fMRI scanning, we also observed that some L2 learners engaged in subvocal reading of materials. As visual information could be transformed into inner language for phonological encoding (Ye & Liu, 2008) and the PreCG (part of middle frontal gyrus, MFG) had been found to respond to verbal working memory tasks

(Fedorenko et al., 2011), particularly in representing and maintaining the complex structures (see also Hickok & Poeppel, 2007), the connection from the pIFG to the PreCG in the present study might be able to assist L2 learners in maintaining complex syntactic representations through inner speech encoding, which aligned with the role of this connection in Chinese L1 syntactic processing as also detected in Chen et al. (2023b).

The activation pattern of the frontoparietal network was also observed in previous AGL studies (Fletcher et al., 1999; Seger et al., 2000). For example, Fletcher et al. (1999) found that the connectivity between the left frontal and parietal regions gradually strengthened during the learning process of finite-state grammar. Seger et al. (2000) also observed the frontoparietal activation patterns during the artificial grammar judgment task. They suggested that this frontoparietal activation pattern was associated with working memory and argued that grammar judgment tasks required the involvement of working memory resources, with the parietal region playing a crucial role in storing and retrieving language information (Smith & Jonides, 1997). In the present study, the negative correlation between LD\_SimS and the contemporaneous PreCG -to- SPL connectivity strength suggested that higher syntactic similarity between Korean (L1) and Chinese (L2) accompanied with lower demands for storing L2 structures, as reflected by the weakened PreCG -to- SPL connectivity strength. This interconnected relationship between the syntactic similarity across languages, the working memory load for syntactic processing and the PreCG -to- SPL connectivity strength reflected the impact of syntactic similarity on L2 processing at the neural level, hence further providing neurobiological evidence supporting the reliance on L2 syntactic processing on the working memory system (McDonald, 2006).

Therefore, we assumed that the frontoparietal connectivity in L2 learners was closely related to the storage and retrieval of L2 syntactic information. Based on the prominent neurobiological models of language processing (Friederici, 2017; Skeide & Friederici, 2016; Xiang et al., 2010), which hypothesized the information flow direction from frontal regions to the posterior regions (including the posterior temporal lobe as well as the parietal regions), we postulated that there might be a frontoparietal pathway wherein pIFG sends the merged constituents to SPL for storage. Previous studies on syntactic processing also consistently validated the role of pIFG as the starting point (e.g., Chen et al., 2021; den Ouden et al., 2012; Xu et al., 2020; Wu et al., 2019). Subsequently, PreCG was assumed to be related to linearize the hierarchical structures generated by the pIFG into syntactic frames or templates as phonological representations (Fedorenko et al., 2011; Hickok and Poeppel 2007; Kaestner et al., 2022; Wheat et al., 2010). And these linear sequences would be sent to SPL for storage. As a memory component, SPL would activate and retrieve the lexical information as well as the linear syntactic templates stored within it for subsequent syntactic processes (Koenigs et al., 2009; Nęcka et al., 2021). Moreover, spatial features/attention and the cognitive control demand might also affect SPL to manipulate the syntactic information. In brief, the pIFG transferred processed syntactic information to the PreCG through internal language encoding. The PreCG linearized hierarchical structures into syntactic frames as phonological representations and transmitted them to the SPL for further storage. The SPL retrieved stored syntactic information and explicit syntactic knowledge from its own storage, transferring them back to the pIFG for subsequent syntactic operations, thus establishing a neural circuit that functionally organizes the dynamic interaction between language and multiple demand networks.

However, as the uSEM approach did not report the site receiving the driving inputs, the possibility of considering SPL as the source region should not be excluded. In this case, participants might extract the stored lexical or constituent information from SPL, and then send this information to pIFG for Merge. The newly merged syntactic objects would then be sent back to SPL for storage, inter-mediated by PreCG.

In summary, the neural circuit underlying Chinese L2 syntactic processing was centered around the SPL as the crucial region potentially for multiple general cognitive functions such as the working memory capacity. In terms of the learner factors, ChiProfS was positively correlated with the strength of SPL activation and LD\_SimS was negatively correlated with the PreCG-SPL connectivity strength, indicating that L2 learners were inclined to rely more on explicit syntactic knowledge and general cognitive abilities during syntactic processing. This neural foundation should back up the complex hierarchical syntactic processing of Chinese L2 learners.

#### 4.3. Differences between the neural correlates of Chinese L1 and L2 syntactic processing

To explore how L2 learners perform in syntactic processing, previous studies often utilized natural materials, yielding mixed findings. Some studies suggested that L2 syntactic processing could reach native-speaker levels (Bowden et al., 2013; Friederici et al., 2002; Grey et al., 2018; Morgan-Short et al., 2010, 2012a, 2012b; Rossi et al., 2006; Steinhauer et al., 2009). However, other research indicated that L2 learners' performance varied with the complexity of syntactic structures, with notable differences emerging primarily during the processing of abstract syntactic representations (Clahsen & Felser, 2018). For example, L2 learners struggled with parsing complex dependency structures, relying on direct lexical associations to establish long-distance filler-gap dependencies rather than employing structure-based gap filling (Clahsen & Felser, 2006a; Clahsen et al., 2010). In this study, we constructed semantically deprived jabberwocky sentences based on authentic Chinese grammatical rules and designed diverse syntactic structures to cover a range of hierarchical embedding depths and dependency lengths. Behavioral results revealed that even highly proficient L2 learners experience difficulties in processing complex hierarchical structures based on Chinese function words. These findings underscore L2 processing limitations at the level of abstract syntactic representation, providing further evidence in support of the SSH.

At the neural level, this study identified distinct activation networks differentiating L2 processing from native language processing. While comparing the L2-specific neural network with the L1 frontotemporal language-general network as reported in our previous studies (Chen et al., 2023b), we found that L2 learners lacked significant activation in the temporal lobe as well as the significant connectivity between the frontal and temporal regions. Previous research had indicated the crucial role of the posterior superior temporal gyrus/superior temporal sulcus (pSTG/STS) in integrating syntactic and semantic information during the processing of complex syntactic structures (Bornkessel et al., 2005), particularly when the syntactic elements were complex or difficult to integrate into the structures (Constable et al., 2004). However, the lack of significant left posterior temporal lobe activation might reflect a deficiency in representing the whole structures with their syntactic categories during processing in Chinese L2 learners,

reflecting the limited processing depth of syntactic representation in L2 learners.

As assumed in SSH, L1 speakers and L2 learners employed different processing strategies for syntactic information (Clahsen & Felser, 2018). L2 learners relied more on the memory system for syntactic comprehension, rather than engaging in syntactic parsing like L1 speakers (Cunnings, 2017a). To be specific, assuming language processing involved cue-based memory retrieval, L2 exhibited less use of abstract syntactic cues when compared with L1. Cunnings (2017b) indicated that the L2 processing relies more on whole-word storage of complex words and less on morphological parsing, resulting in shallower morphological processing in the L2 than in the L1. Previous ERP studies provided neural evidence supporting this distinction, showing that L2 learners often exhibited response patterns to syntactic violations that differed from those of native speakers (Guo et al., 2009; Hahne, 2001; Hahne & Friederici, 2001; Pakulak & Neville, 2011; Zhang, 2019). For instance, highly proficient L2 learners displayed delayed or absent LAN effects, indicating weaker automatic syntactic processing (Hahne & Friederici, 2001; Weber-Fox & Neville, 1996). Even when L2 learners transferred certain syntactic rules from their native language, they still failed to achieve native-like automatic processing of syntactic relations which was reflected in P600 that were often diminished in amplitude and delayed in latency (Hahne, 2001; Pakulak & Neville, 2011). In this study, we found that L2 processing differed from L1 processing, particularly in the lack of activation in the pTL responsible for the integration of complex structures. Instead, L2 processing relied more on the SPL, which was associated with memory storage. This reliance on memory systems highlighted a distinct processing mechanism for L2, providing spatially neural evidence aligning with the SSH. It is also worth noting that these previous studies found shallower processing of syntactic cues mainly in morphologically rich languages. As Chinese lacked morphological information changes, syntactic processing based on Chinese function words advanced the development of L2 processing theories, extended findings originally derived from Indo-European languages and emphasized the shared characteristics of L2 processing.

In addition to the prominent SSH we began with in the present study, the findings of this study could also be extended to support other important models/hypotheses in the field of L2 acquisition, which are complementary rather than contradictory. For instance, our current finding could provide further evidence to Ullman's declarative/procedural (DP) model. According to Ullman (2001a, 2001b, 2006, 2016, 2020), language processing primarily relied on two distinct memory systems: the declarative memory system, which handled idiosyncratic knowledge such as vocabulary, irregular morphological forms, prefabricated phrases and idioms, and the procedural memory system, responsible for rule-based combinations like regular morphology and structural processing. While both native speakers and L2 learners utilized these memory systems, L2 learners often relied more heavily on the declarative memory system to compensate for limited automatic processing abilities (Ullman, 2001b). This strategy allowed them to process structures that would typically be handled by the procedural memory system, as retrieving information directly from declarative memory was safer, more efficient, and less effortful. Accordingly, a neuroanatomical meta-analysis of grammatical learning reported activations in the IFG and the SPL (Tagarelli et al., 2019). As the IFG was linked to declarative memory (Ullman, 2004, 2016) and the SPL was associated with encoding and retrieving information stored in declarative memory (Wagner et al., 2005), these activations aligned

with the declarative memory system's involvement in L2 grammar processing, as proposed in the DP model. In this study, significant activations of the pIFG and SPL in L2 learners were observed. The activation of this frontoparietal network highlighted L2 learners' greater reliance on the declarative memory system during online syntactic processing, thus providing robust evidence to support the DP model from Chinese L2 perspective.

Therefore, the recruitment of the frontoparietal network which supported working memory maintenance and other related general cognitive capacities, instead of the general frontotemporal network for language (esp., syntactic) processing, provided neural evidence for the reliance of L2 syntactic processing on more strategies, thus in support of the SSH. The present findings could also provide further evidence to other related theories of L2 processing, such as the DP model, in a complementary fashion.

## 5. Limitations

Since the known definitive measurements of language distance were calculated primarily on the features of words (Cargnelli et al., 2022), they were not suitable for measuring the bilingual syntactic structure similarity particularly at the above-lexicon level. This study employed a Likert 5-point scale to quantify linguistic similarity by asking Korean participants to rate their subjective perceptions of how similar the experimental Chinese L2 syntactic structures were to the corresponding structures in Korean. These ratings of psychological distance were used as quantitative data for linguistic similarity. Although this method is inevitably subjective and indirect, the psychological distance toward different structures varies among individuals, making it necessary to adopt a language-experience perspective and incorporate such bilingual experiences as regressors in the statistical design (Hell, 2023). Here we deemed the "linguistic similarity" as a potential confounder and regressed out it as a covariate to further purify the imaging results in a relatively conservative way. Future research could consider utilizing large language models to compute structure similarity, providing a more precise quantification of the influence of L2 factors.

Building on these methodological considerations, this study further explored the effective connectivity network underlying syntactic processing in L2 learners. Given that the uSEM approach did not report the site receiving the driving inputs, it was challenging to determine the starting region within the effective connectivity network. However, based on neurobiological models (Skeide & Friederici, 2016; Xiang et al., 2010) and neuroimaging studies investigating syntactic merge processing (Chen et al., 2021; den Ouden et al., 2012; Xu et al., 2020; Wu et al., 2019), this study primarily identified the pIFG as the starting point of the frontoparietal network. At the same time, we acknowledged the alternative possibility of information flow originating in the parietal regions and progressing from posterior to frontal areas, further highlighting the complexity of syntactic processing in L2 learners.

As a result of these investigations, a distinct frontoparietal network for Chinese L2 at the neural level was identified, which differed from the frontotemporal network observed in previous studies on Chinese native syntactic processing (Chen et al., 2023b), highlighting the unique characteristics of L2. To deepen our understanding of these distinctions, future research could focus on L2 learners with varying proficiency levels or adopt longitudinal designs to track neural changes over time, thereby scrutinizing systematic neural changes in syntactic processing and exploring how specific learner variables influence language acquisition (Hell, 2023).

## 6. Conclusion

This fMRI study employed the jabberwocky sentence processing paradigm based on real Chinese function words to investigate the neural basis of Chinese L2 syntactic processing. The results showed that, similar to Chinese L1 native speakers, the frontal regions including pIFG and PreCG in the language network were consistently activated, while the activation of SPL and the frontoparietal network revealed the interaction between language and multiple demand networks for fostering Chinese L2 syntactic processing. By comparing the behavioral performance and neural mechanisms between Chinese L2 learners and L1 speakers, this study provided reliable neurobiological evidence from the spatial dimension for the SSH as well as other related models especially by considering the case of Chinese L2 syntactic processing.

**Supplementary material.** The supplementary material for this article can be found at <http://doi.org/10.1017/S1366728925100138>.

**Data availability statement.** Anonymized data will be made available upon reasonable requests and collaborative agreement addressed to the coauthors. Researchers wishing to obtain the data must contact the Max Planck Partner Group, College of Chinese Language and Culture, Beijing Normal University, to sign a formal data sharing agreement. The Matlab toolboxes used for the imaging data analyses are freely available online: (a) DPARSF 5.1 Advanced Edition: <http://www.rfmri.org/DPARSF>; (b) SPM12: <https://www.fil.ion.ucl.ac.uk/spm/> and (c) GIMME: <https://gimme.web.unc.edu/>.

**Acknowledgements.** The authors wish to thank the two anonymous reviewers for their insightful and valuable comments and all participants who took part in this study. Special thanks are extended to Yang Liu, Siying Lin, Keyi Kang, Luping Wang, Yi Song and Lu Li for their support in the behavioral and fMRI data collection. The authors also thank Junjie Wu, Siyuan Zhou, Xingfang Qu and other colleagues for their constructive input.

**Author contribution.** C.G. and J.G. (co-first-authors): Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, and Writing – original draft. L.Z.: Data curation, Validation, and Writing – review and editing. Z.L.: Data curation and Writing – review and editing. L.C. and L.F. (co-corresponding-authors): Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Supervision, Validation, and Writing – review and editing.

**Funding statement.** This work was supported by the 2022 International Chinese Language Education Research Project (No. 22YH08A) from the Center for Language Education and Cooperation, Ministry of Education of the People's Republic of China and by the National Social Science Foundation of China (No. 22CYY017). It was also funded by the STI 2030—Major Projects + 2021ZD0200500 and the National Key R&D Program of China (2019YFA0709503).

**Competing interests.** The authors declare no competing interests.

## References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Bernal, B., & Ardila, A. (2009). The role of the arcuate fasciculus in conduction aphasia. *Brain*, *132*(9), 2309–2316. <https://doi.org/10.1093/brain/awp206>.
- Blank, I. A., & Fedorenko, E. (2017). Domain-general brain regions do not track linguistic input as closely as language-selective regions. *The Journal of Neuroscience*, *37*(41), 9999–10011. <https://doi.org/10.1523/JNEUROSCI.3642-16.2017>.
- 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>.
- Bowden, H. W., Steinhauer, K., Sanz, C., & Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia*, *51*(13), 2492–2511. <https://doi.org/10.1016/j.neuropsychologia.2013.09.004>.
- Cargnelutti, E., Tomasino, B., & Fabbro, F. (2022). Effects of linguistic distance on second language brain activations in bilinguals: An exploratory coordinate-based meta-analysis. *Frontiers in Human Neuroscience*, *15*, 744489. <https://doi.org/10.3389/fnhum.2021.744489>.
- Chen, L., Gao, C., Li, Z., Zaccarella, E., Friederici, A. D., & Feng, L. (2023b). Frontotemporal effective connectivity revealed a language-general syntactic network for Mandarin Chinese. *Journal of Neurolinguistics*, *66*, 101127–101127. <https://doi.org/10.1016/j.jneuroling.2023.101127>.
- Chen, L., Goucha, T., Männel, C., Friederici, A. D., & Zaccarella, E. (2021). Hierarchical syntactic processing is beyond mere associating: Functional magnetic resonance imaging evidence from a novel artificial grammar. *Human Brain Mapping*, *42*(10), 3253–3268. <https://doi.org/10.1002/hbm.25432>.
- Chen, L., Yang, M., Gao, F., Fang, Z., Wang, P., & Feng, L. (2023a). Mandarin Chinese L1 and L2 complex sentence reading reveals a consistent electrophysiological pattern of highly interactive syntactic and semantic processing: An ERP study. *Frontiers in Psychology*, *14*. <https://doi.org/10.3389/fpsyg.2023.1143062>.
- Chomsky, N. (1995). *The minimalist program*. The MIT Press.
- Clahsen, H., & Felser, C. (2006a). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, *27*(1), 107–126. <https://doi.org/10.1017/S0142716406060206>.
- Clahsen, H., & Felser, C. (2006b). How native-like is non-native language processing? *Trends in Cognitive Sciences*, *10*(12), 564–570. <https://doi.org/10.1016/j.tics.2006.10.002>.
- Clahsen, H., & Felser, C. (2006c). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(1), 3–42. <https://doi.org/10.1017/S0142716406060024>.
- Clahsen, H., & Felser, C. (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, *40*(3), 693–706. <https://doi.org/10.1017/S0272263117000250>.
- Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning*, *60*, 21–43. <https://doi.org/10.1111/j.1467-9922.2009.00550.x>.
- Constable, R. T., Pugh, K. R., Berroya, E., Mencl, W. E., Westerveld, M., Ni, W., et al. (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>.
- Cummings, I. (2017a). Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*, *20*(4), 659–678. <https://doi.org/10.1017/s1366728916000675>.
- Cummings, I. (2017b). Interference in native and non-native sentence processing. *Bilingualism: Language and Cognition*, *20*(4), 712–721. <https://doi.org/10.1017/S1366728916001243>.
- de Vries, M. H., Barth, A. C. R., Maiworm, S., Knecht, S., Zwitterlood, P., & Flöel, A. (2010). Electrical stimulation of Broca's area enhances implicit learning of an artificial grammar. *Journal of Cognitive Neuroscience*, *22*(11), 2427–2436. <https://doi.org/10.1162/jocn.2009.21385>.
- den Ouden, D.-B., Saur, D., Mader, W., Schelter, B., Lukic, S., Wali, E., Timmer, J., & 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>.
- Diachek, E., Blank, I., Siegelman, M., Affourtit, J., & Fedorenko, E. (2020). The domain-general multiple demand (MD) network does not support core aspects of language comprehension: A large-scale fMRI investigation. *The Journal of Neuroscience*, *40*(23), 4536–4550. <https://doi.org/10.1523/JNEUROSCI.2036-19.2020>.
- Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension. *Frontiers in Psychology*, *5*, 335–335. <https://doi.org/10.3389/fpsyg.2014.00335>.
- Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences – PNAS*, *108*(39), 16428–16433. <https://doi.org/10.1073/pnas.1112937108>.

- Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: Defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, 104(2), 1177–1194. <https://doi.org/10.1152/jn.00032.2010>.
- Feng, L., Feng, H., Bai, S., & Wu, J. (2020). An analysis of proficiency test for CSL (Chinese as second language) based on fixed-ratio cloze questions. *Applied Linguistics*, 3, 69–79. <https://doi.org/10.16499/j.cnki.1003-5397.2020.03.011>.
- Fletcher, P., Büchel, C., Josephs, O., Friston, K., & Dolan, R. (1999). Learning-related neuronal responses in prefrontal cortex studied with functional neuroimaging. *Cerebral Cortex*, 9(2), 168–178. <https://doi.org/10.1093/cercor/9.2.168>.
- Friederici, A. D. (2017). *Language in our brain: The origins of a uniquely human capacity*. The MIT Press.
- Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences – PNAS*, 99(1), 529–534. <https://doi.org/10.1073/pnas.012611199>.
- Gates, K. M., Molenaar, P. C. M., Hillary, F. G., & Slobounov, S. (2011). Extended unified SEM approach for modeling event-related fMRI data. *NeuroImage*, 54(2), 1151–1158. <https://doi.org/10.1016/j.neuroimage.2010.08.051>.
- Geva, S., Schneider, L. M., Khan, S., Lorca-Puls, D. L., Gajardo-Vidal, A., PLORAS team, Hope, T. M. H., Green, D. W., & Price, C. J. (2023). Enhanced left superior parietal activation during successful speech production in patients with left dorsal striatal damage and error-prone neurotypical participants. *Cerebral Cortex*, 33(7), 3437–3453. <https://doi.org/10.1093/cercor/bhac282>.
- Greenberg, A. S., Esterman, M., Wilson, D., Serences, J. T., & Yantis, S. (2010). Control of spatial and feature-based attention in frontoparietal cortex. *The Journal of Neuroscience*, 30(43), 14330–14339. <https://doi.org/10.1523/jneurosci.4248-09.2010>.
- Grey, S., Sanz, C., Morgan-Short, K., & Ullman, M. T. (2018). Bilingual and monolingual adults learning an additional language: ERPs reveal differences in syntactic processing. *Bilingualism*, 21(5), 970–994. <https://doi.org/10.1017/S1366728917000426>.
- Guo, J., Guo, T., Yan, Y., Jiang, N., & Peng, D. (2009). ERP evidence for different strategies employed by native speakers and L2 learners in sentence processing. *Journal of Neurolinguistics*, 22(2), 123–134. <https://doi.org/10.1016/j.jneuroling.2008.09.001>.
- Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, 30(3), 251–266. <https://doi.org/10.1023/a:1010490917575>.
- Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism*, 4(2), 123–141. <https://doi.org/10.1017/S1366728901000232>.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The Faculty of Language: What is it, who has it, and how did it evolve? *Science*, 298(5598), 1569–1579. <https://doi.org/10.1126/science.298.5598.1569>.
- Hell, J. G. (2023). The neurocognitive underpinnings of second language processing: Knowledge gains from the past and future outlook. *Language Learning*, 73(S2), 95–138. <https://doi.org/10.1111/lang.12601>.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, 8(5), 393–402. <https://doi.org/10.1038/nrn2113>.
- Hou, Z., Li, H., Gao, L., Ou, J., & Xu, M. (2024). Differential neural representations of syntactic and semantic information across languages in Chinese-English bilinguals. *NeuroImage*, 303, 120928. <https://doi.org/10.1016/j.neuroimage.2024.120928>.
- Huang, B., Liao, X. (2011). *Modern Chinese* (updated 5th editions) (现代汉语). Higher Education Press.
- Iwabuchi, T., & Makuuchi, M. (2021). When a sentence loses semantics: Selective involvement of a left anterior temporal subregion in semantic processing. *The European Journal of Neuroscience*, 53(3), 929–942. <https://doi.org/10.1111/ejn.15022>.
- Kaestner, E., Wu, X., Friedman, D., Dugan, P., Devinsky, O., Carlson, C., Doyle, W., Thesen, T., & Halgren, E. (2022). The Precentral Gyrus contributions to the early time-course of grapheme-to-phoneme conversion. *Neurobiology of Language*, 3(1), 18–45. [https://doi.org/10.1162/nol\\_a\\_00047](https://doi.org/10.1162/nol_a_00047).
- Kirkici, B., & Clahsen, H. (2013). Inflection and derivation in native and non-native language processing: Masked priming experiments on Turkish. *Bilingualism: Language and Cognition*, 16(4), 776–791. <https://doi.org/10.1017/S1366728912000648>.
- Koenigs, M., Barbey, A. K., Postle, B. R., & Grafman, J. (2009). Superior parietal cortex is critical for the manipulation of information in working memory. *The Journal of Neuroscience*, 29(47), 14980–14986. <https://doi.org/10.1523/JNEUROSCI.3706-09.2009>.
- Kovelman, I., Baker, S. A., & Petitto, L. A. (2008). Bilingual and monolingual brains compared: A functional magnetic resonance imaging investigation of syntactic processing and a possible “neural signature” of bilingualism. *Journal of Cognitive Neuroscience*, 20(1), 153–169. <https://doi.org/10.1162/jocn.2008.20011>.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>.
- Lee, H., Devlin, J. T., Shakeshaft, C., Stewart, L. H., Brennan, A., Glensman, J., Pitcher, K., Crinion, J., Mechelli, A., Frackowiak, R. S. J., Green, D. W., & Price, C. J. (2007). Anatomical traces of vocabulary acquisition in the adolescent brain. *The Journal of Neuroscience*, 27(5), 1184–1189. <https://doi.org/10.1523/JNEUROSCI.4442-06.2007>.
- Lobier, M., Peyrin, C., Le Bas, J.-F., & Valdois, S. (2012). Pre-orthographic character string processing and parietal cortex: A role for visual attention in reading? *Neuropsychologia*, 50(9), 2195–2204. <https://doi.org/10.1016/j.neuropsychologia.2012.05.023>.
- MacGregor, L. J., Gilbert, R. A., Balewski, Z., Mitchell, D. J., Erzincioğlu, S. W., Rodd, J. M., Duncan, J., Fedorenko, E., & Davis, M. H. (2022). Causal contributions of the domain-general (multiple-demand) and the language-selective brain networks to perceptual and semantic challenges in speech comprehension. *Neurobiology of Language*, 3(4), 665–698. [https://doi.org/10.1162/nol\\_a\\_00081](https://doi.org/10.1162/nol_a_00081).
- Maran, M., Friederici, A. D., & Zaccarella, E. (2022). Syntax through the looking glass: A review on two-word linguistic processing across behavioral, neuroimaging and neurostimulation studies. *Neuroscience and Biobehavioral Reviews*, 142. <https://doi.org/10.1016/j.neubiorev.2022.104881>.
- 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>.
- McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55(3), 381–401. <https://doi.org/10.1016/j.jml.2006.06.006>.
- Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Neurolinguistics structural plasticity in the bilingual brain. *Nature*, 431(7010), 757–757. <https://doi.org/10.1038/431757a>.
- Ministry of Education of the People's Republic of China. (2021). Chinese proficiency grading standards for international Chinese language education (GF0025-2021).
- Morgan-Short, K., Finger, I., Grey, S., Ullman, M. T., & Stamatakis, E. A. (2012a). Second language processing shows increased native-like neural responses after months of no exposure. *PLoS One*, 7(3), e32974–e32974. <https://doi.org/10.1371/journal.pone.0032974>.
- Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning*, 60(1), 154–193. <https://doi.org/10.1111/j.1467-9922.2009.00554.x>.
- Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012b). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, 24(4), 933–947. [https://doi.org/10.1162/jocn\\_a\\_00119](https://doi.org/10.1162/jocn_a_00119).
- Nęcka, E., Gruszka, A., Hampshire, A., Sarzyńska-Wawer, J., Anicai, A. E., Orzechowski, J., Nowak, M., Wójcik, N., Sandrone, S., & Soreq, E. (2021). The effects of working memory training on brain activity. *Brain Sciences*, 11(2), 155. <https://doi.org/10.3390/brainsci11020155>.
- Ohta, S., Fukui, N., & Sakai, K. L. (2013). Syntactic computation in the human brain: The degree of merger as a key factor. *PLoS One*, 8(2), e56230–e56230. <https://doi.org/10.1371/journal.pone.0056230>.

- Osaka, N., Osaka, M., Kondo, H., Morishita, M., Fukuyama, H., & Shibasaki, H. (2004). The neural basis of executive function in working memory: An fMRI study based on individual differences. *NeuroImage*, 21(2), 623–631. <https://doi.org/10.1016/j.neuroimage.2003.09.069>.
- Pakulak, E., & Neville, H. J. (2011). Maturational constraints on the recruitment of early processes for syntactic processing. *Journal of Cognitive Neuroscience*, 23(10), 2752–2765. <https://doi.org/10.1162/jocn.2010.21586>.
- Pallier, C., Devauchelle, A. D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences*, 108(6), 2522, 2527. <https://doi.org/10.1073/pnas.1018711108>.
- Ritz, H., & Shenhav, A. (2024). Orthogonal neural encoding of targets and distractors supports multivariate cognitive control. *Nature Human Behaviour*, 8(5), 945–961. <https://doi.org/10.1038/s41562-024-01826-7>.
- Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, 18(12), 2030–2048. <https://doi.org/10.1162/jocn.2006.18.12.2030>.
- Sakai, K. L., Miura, K., Narafu, N., & Muraishi, Y. (2004). Correlated functional changes of the prefrontal cortex in twins induced by classroom education of second language. *Cerebral Cortex*, 14(11), 1233–1239. <https://doi.org/10.1093/cercor/bhh084>.
- Scherer, L. C., Fonseca, R. P., Amiri, M., Adrover-Roig, D., Marcotte, K., Giroux, F., Senhadji, N., Benali, H., Lesage, F., & Ansaldi, A. I. (2012). Syntactic processing in bilinguals: An fNIRS study. *Brain and Language*, 121(2), 144–151. <https://doi.org/10.1016/j.bandl.2011.09.009>.
- Seger, C. A., Prabhakaran, V., & Poldrack, R. A. (2000). Neural activity differs between explicit and implicit learning of artificial grammar strings: An fMRI study. *Psychobiology*, 28(3), 283–292. <https://doi.org/10.3758/BF03331987>.
- Shomstein, S. (2012). Cognitive functions of the posterior parietal cortex: Top-down and bottom-up attentional control. *Frontiers in Integrative Neuroscience*, 6, 38–38. <https://doi.org/10.3389/fnint.2012.00038>.
- Skeide, M. A., & Friederici, A. D. (2016). The ontogeny of the cortical language network. *Nature Reviews. Neuroscience*, 17(5), 323–332. <https://doi.org/10.1038/nrn.2016.23>.
- Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology*, 33(1), 5–42. <https://doi.org/10.1006/cogp.1997.0658>.
- Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research*, 25(1), 13–41. <https://doi.org/10.1177/0267658308098995>.
- Suh, S., Yoon, H. W., Lee, S., Chung, J.-Y., Cho, Z.-H., & Park, H. (2007). Effects of syntactic complexity in L1 and L2: an fMRI study of Korean–English bilinguals. *Brain Research*, 1136, 178–189. <https://doi.org/10.1016/j.brainres.2006.12.043>.
- Sulpizio, V., Fattori, P., Pitzalis, S., & Galletti, C. (2023). Functional organization of the caudal part of the human superior parietal lobule. *Neuroscience and Biobehavioral Reviews*, 153, 105357–105357. <https://doi.org/10.1016/j.neubiorev.2023.105357>.
- Sun, Y., Yang, Y., Desroches, A. S., Liu, L., & Peng, D. (2011). The role of the ventral and dorsal pathways in reading Chinese characters and English words. *Brain and Language*, 119(2), 80–88. <https://doi.org/10.1016/j.bandl.2011.03.012>.
- Szczepanski, S. M., Konen, C. S., & Kastner, S. (2010). Mechanisms of spatial attention control in frontal and parietal cortex. *The Journal of Neuroscience*, 30(1), 148–160. <https://doi.org/10.1523/JNEUROSCI.3862-09.2010>.
- Tagarelli, K. M., Shattuck, K. F., Turkeltaub, P. E., & Ullman, M. T. (2019). Language learning in the adult brain: A neuroanatomical meta-analysis of lexical and grammatical learning. *NeuroImage*, 193, 178–200. <https://doi.org/10.1016/j.neuroimage.2019.02.061>.
- Tang, SW. (2019). *Formal Chinese syntax* (2nd ed.) (形式汉语语法学). Shanghai Educational Publishing House.
- Ullman, M. (2016). The declarative/procedural model: A neurobiological model of language learning, knowledge, and use. In G. Hickock & S. L. Small (Eds.), *Neurobiology of language* (pp. 953–968). New York: Academic.
- Ullman, M. (2020). The declarative/procedural model: A neurobiologically motivated theory of first and second language. In B. van Patten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 128–161). London: Routledge.
- Ullman, M. T. (2001a). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews. Neuroscience*, 2(10), 717–726. <https://doi.org/10.1038/35094573>.
- Ullman, M. T. (2001b). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism*, 4(2), 105–122. <https://doi.org/10.1017/S1366728901000220>.
- Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92(1), 231–270. <https://doi.org/10.1016/j.cognition.2003.10.008>.
- Ullman, M. T. (2006). The declarative/procedural model and the shallow structure hypothesis. *Applied Psycholinguistics*, 27(1), 97–105. <https://doi.org/10.1017/S014271640606019X>.
- Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences*, 9(9), 445–453. <https://doi.org/10.1016/j.tics.2005.07.001>.
- Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8(3), 231–256. <https://doi.org/10.1162/jocn.1996.8.3.231>.
- Wheat, K. L., Cornelissen, P. L., Frost, S. J., & Hansen, P. C. (2010). During visual word recognition, phonology is accessed within 100 ms and may be mediated by a speech production code: Evidence from magnetoencephalography. *The Journal of Neuroscience*, 30(15), 5229–5233. <https://doi.org/10.1523/JNEUROSCI.4448-09.2010>.
- Wu, C., Zaccarella, E., & Friederici, A. D. (2019). Universal neural basis of structure building evidenced by network modulations emerging from Broca's area: The case of Chinese. *Human Brain Mapping*, 40(6), 1705–1717. <https://doi.org/10.1002/hbm.24482>.
- Wu, C. Y., Ho, M. H., & Chen, S. H. (2012). A meta-analysis of fMRI studies on Chinese orthographic, phonological, and semantic processing. *NeuroImage*, 63(1), 381–391. <https://doi.org/10.1016/j.neuroimage.2012.06.047>.
- Xiang, H., Dediu, D., Roberts, L., Oort, E. v., Norris, D. G., & Hagoort, P. (2012). The structural connectivity underpinning language aptitude, working memory, and IQ in the Perisylvian language network. *Language Learning*, 62(s2), 110–130. <https://doi.org/10.1111/j.1467-9922.2012.00708.x>.
- Xiang, H.-D., Fonteijn, H. M., Norris, D. G., & Hagoort, P. (2010). Topographical functional connectivity pattern in the Perisylvian language networks. *Cerebral Cortex*, 20(3), 549–560. <https://doi.org/10.1093/cercor/bhp119>.
- Xu, K., Wu, D. H., & Duann, J. R. (2020). Dynamic brain connectivity attuned to the complexity of relative clause sentences revealed by a single-trial analysis. *NeuroImage*, 217, 116920. <https://doi.org/10.1016/j.neuroimage.2020.116920>.
- Yan, C.-G., Wang, X.-D., Zuo, X.-N., & Zang, Y.-F. (2016). DPABI: Data processing & analysis for (resting-state) brain imaging. *Neuroinformatics*, 14(3), 339–351. <https://doi.org/10.1007/s12021-016-9299-4>.
- Yang, J., & Li, P. (2012). Brain networks of explicit and implicit learning. *PLoS One*, 7(8), e42993–e42993. <https://doi.org/10.1371/journal.pone.0042993>.
- Ye, M., & Liu, X. (2008). A study on the modality effect in true memory and false memory. *Psychological Science*, 05, 1104–1107. <https://doi.org/10.16719/j.cnki.1671-6981.2008.05.016>.
- 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>.
- Zhang, H. (2019). *A contrastive study of the Chinese double-object construction coercion mechanism between Chinese native speakers and Chinese second language learners*. (Doctoral dissertation). Beijing Normal University.
- Zhao, Y. (2018). Forty years of second language acquisition studies of Chinese. *International Chinese Language Education*, 04, 92–101. <https://doi.org/CNKI:SUN:GZZY.0.2018-04-010>.