Bilingualism: Language and Cognition

cambridge.org/bil

The stability and instability of the language control network: A longitudinal resting-state functional magnetic resonance imaging study

Zilong Li[1], Cong Liu[2], Xin Pan[1], Guosheng Ding[3] and Ruiming Wang[1]

[1]Key Laboratory of Brain, Cognition and Education Sciences, & Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents, Ministry of Education, and Center for Studies of Psychological Application, School of Psychology, South China Normal University, Guangzhou, P.R. China; [2]Brain, Cognition, and Language Learning Laboratory, Department of Psychology, School of Education Science, Qingdao University, Qingdao, P. R. China and [3]State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Faculty of Psychology, Beijing Normal University, Beijing, P. R. China

Research Article

Cite this article: Li, Z., Liu, C., Pan, X., Ding, G. and Wang, R. (2026). The stability and instability of the language control network: A longitudinal resting-state functional magnetic resonance imaging study. *Bilingualism: Language and Cognition* **29**, 1–13. https://doi.org/10.1017/S1366728925000185

Received: 21 February 2024
Revised: 25 December 2024
Accepted: 3 February 2025
First published online: 28 February 2025

Keywords:
resting-state functional connectivity; connectome fingerprint identification; stability; language control; temporal complexity

Corresponding authors:
Ruiming Wang and Guosheng Ding;
Emails: wangrm@scnu.edu.cn; dinggsh@bnu.edu.cn

Z.L. and C.L. contributed equally to this work.

This article has earned badges for transparent research practices: Open Data and Open Materials. For details see the Data Availability Statement.

Abstract

This study investigates the stability and instability of the language control network in bilinguals using longitudinal resting-state functional magnetic resonance imaging (rs-fMRI) data. We compared the language control network of Chinese university students majoring in English with those not, using three other functional networks as controls. Results indicate that the English major group exhibits reduced stability and increased instability in the language control network compared with the non-English major group. This suggests that second language (L2) learning experience may induce adaptive neural changes. Moreover, the coexistence of stability and instability in the language control network appears less modular in the English major group, implying a more integrated response to language experience. Notably, these results were not observed in the control networks. Overall, these findings enhance the understanding of bilingual language control and the impact of L2 learning on neural plasticity.

Highlights

- This study focused on both the stability and instability of language control network.
- English majors showed reduced stability and increased instability in the network.
- Coexistence of stability and instability was less modular in English majors.
- Results deepen understanding of bilingual language control and neural plasticity.

1. Introduction

In multicultural and multilingual societies, bilinguals are required to switch between languages. This process requires bilingual language control, which refers to the cognitive abilities that minimize interference from the non-target language when bilinguals use the target language (Declerck et al., 2015; Jiao et al., 2025a, b). Previous studies have proposed that language control processes involve a complex neural network that includes multiple cortical and subcortical brain regions (Abutalebi & Green, 2016; Green & Abutalebi, 2013; Jiao et al., 2022; Liu et al., 2021a, b, c; Wang et al., 2016; Wu et al., 2019). A meta-analysis of studies on bilingual language control identified eight brain regions as part of the language control network: the pre-supplementary motor area (pre-SMA), left inferior frontal gyrus (IFG), left middle temporal gyrus (MFG), left middle frontal gyrus, right precentral gyrus (PrCG), right superior temporal gyrus (STG) and bilateral caudate (Luk et al., 2012). In addition to these regions, the dorsal anterior cingulate cortex (dACC) and the left inferior parietal lobe (IPL) is also considered important parts of the language control network (Abutalebi et al., 2013; Abutalebi & Green, 2016; Barbeau et al., 2017; Branzi et al., 2016; Liu et al., 2021a).

Previous studies have shown that resting-state brain connectomes are stable (Amico & Goñi, 2018; Bari et al., 2019; Cao et al., 2024; Chen et al., 2015; Finn et al., 2015; Kaufmann et al., 2017; Liu et al., 2018; Munsell et al., 2020; Ravindra et al., 2019; Yeh et al., 2016; Zhang et al., 2022; Zuo & Xing, 2014). This stability is peculiar, similar to that of a fingerprint, indicating that the brain connectome is highly variable across individuals (unique) but highly reliable (stable) at different times in each individual (Dufford et al., 2021). Notably, a high degree of stability is present in the prefrontal lobe (Finn et al., 2015; Liu et al., 2018; Liu et al., 2020; Yeh et al., 2016; Zhang et al., 2022). Furthermore, several brain regions in the language control network belong to the frontoparietal lobe. Therefore, stability may also exist in the language control network. Notably, many studies have also indicated the stability of the language control network. For example,

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Liu et al. (2021a) found that long-term second language (L2) learning modified only the functional connectivity strength between two brain regions of the language control network, whereas the other five connections remained unchanged. Furthermore, the extended classroom instruction did not modify the grey matter volume of the left anterior cingulate cortex or caudate, which are components of the language control network (Liu et al., 2021c). Therefore, the language control network may be inferred to have a certain degree of stability; however, this has not been directly tested.

In addition to stability, the brain also demonstrates instability. Dufford et al. (2021) revealed that the brains of 1-year-old infants are not as stable as those of adults and proposed that this instability is due to extensive and significant alterations in the infant's brain. Similarly, the language control network is believed to undergo significant changes during adulthood (Antoniou, 2019). The representative theories about the instability of language control networks are the dynamic restructuring model (DRM) (Pliatsikas, 2020) and the adaptive control hypothesis (ACH) (Green & Abutalebi, 2013). DRM suggests that language learning and switching experiences lead to dynamic restructuring. The ACH suggests that the language control network adapts to the L2 learning experience (Abutalebi & Green, 2016; Green & Abutalebi, 2013). Both DRM and ACH propose that language experience shapes the language control network. This proposal is supported by the findings that the connectome of language control networks is significantly changed after long-term L2 learning (Barbeau et al., 2017; Kang et al., 2017) and that both regional brain activity and functional connectivity are sensitive to L2 experience (Abutalebi & Rietbergen, 2014; Alotaibi et al., 2023; Baum & Titone, 2014; Bialystok, 2014; Kroll & Chiarello, 2016; Li et al., 2014; Pliatsikas, 2020; Sander et al., 2023; Tu et al., 2015; Zhang et al., 2023; Zou et al., 2012).

In conclusion, it can be stated that language experience, particularly L2 experience, exerts an influence on the language control network. However, this impact does not result in a comprehensive alteration of the language control network, nor does it lead to substantial modifications in all its constituent brain regions and connections. The findings of both Liu et al. (2021a) and Liu et al. (2021c) indicate that after 1-year's L2 learning, only a minor proportion of the language control network showed alterations, while the majority remained unaffected. For example, in Liu et al. (2021a), with respect to the four regions in the language control network, only one demonstrates significant change, while the other three remain stable. DRM also indicates that only some brain regions and white matter connections are susceptible to change at each stage, while others remain unaltered. This may be due to the fact that the effects of L2 experience are not disseminated to all connections and brain regions, or it may be that specific connections and brain regions are more resistant to change. Nevertheless, it can be concluded that only one part of the language control network is influenced by L2 experience, while the others remain unaffected.

The language control network may be simultaneously stable and unstable, and it may seem contradictory to think that a brain network has two contrasting properties. However, Spear (2013) proposed that this is possible because there is a balance between stability and instability in the brain. Specifically, instability indicates that the brain can adapt to new experiences throughout life, whereas stability indicates that the brain can resist certain changes to maintain reliable cognition and behavioural patterns. The same is true for the language control network, which requires a stable structure to fulfil this function of language control, but also has instability to accommodate bilingual experiences. Since stability and instability can coexist in the language control network, understanding how they coexist is essential. Fedorenko and Thompson-Schill (2014) provide valuable insights into two potential coexistence models. The first proposes that distinct functions are achieved through the combination of a core module and specific peripheral modules. Each module comprises various nodes (different brain regions). In the language control network, certain brain regions are involved in certain functions that lead to perceptual instability, whereas the others are responsible for maintaining stability in language control abilities. However, the second model suggests no modularity in the language network; the entire language network performs its corresponding functions by reorganizing itself into different patterns. Brain regions within the language control network adapt to the environment via a specific pattern of connectivity. This pattern must exhibit significant instability, whereas the other patterns must remain stable to fulfil certain requirements.

Therefore, we propose two alternative hypotheses regarding the coexistence of stability and instability in the language control network. The first is that the coexistence pattern is modular, with one set of nodes supporting stability and another set supporting instability (see Figure 1A). This hypothesis suggests tight inner connections/edges between nodes within the same set but loose connections between nodes from different sets. In this case, the nodes of the language control network can be modularized into two clusters, each focusing on a specific feature. The other hypothesis is that coexistence may not be modular; the entire network implements language control through a specific connection pattern, whereas no other connection pattern is involved in language control. This implies that some connections in the language control network may easily change over time, whereas others remain stable. In this case, the instability and stability of the network are embodied by separate connections/edges (see Figure 1A).

In addition, the coexistence pattern of the language control network may be modulated by L2 experience. Different levels of L2 experience can lead to different levels of brain remodelling (Caffarra et al., 2015; Cargnelutti et al., 2022; DeLuca et al., 2024; Korenar et al., 2023; Kotz, 2009; Połczyńska & Bookheimer, 2021; Wei et al., 2024). The DRM also suggests that with an increase in L2 experience, instability manifests in three distinct patterns. Notably, diverse levels of language proficiency elicit varying impacts on the language control network. L2 experience might influence the balance between stability and instability in the language control network. An increased bilingual experience may lead to a less stable network. However, the morphological appearance of the pattern of higher instability coexisting with lower stability has not yet been adequately investigated.

In this study, we aim to investigate three primary questions. First, to what extent does the language control network exhibit stability and instability? Second, if the language control network manifests both stability and instability, how do these two properties coexist within a single network? Third, does bilingual experience influence the degree of stability and instability, as well as the coexistence patterns of these properties? To address these questions, we utilized resting-state functional magnetic resonance imaging (fMRI) data collected at 1-year intervals. These datasets are scanned from native-speaking Chinese college students, with one group majoring in English (exposed to extensive, varied classroom L2 learning) and another group majoring in non-English (receiving less frequent and diverse classroom L2 exposure). The degree of stability and instability within the language control network were quantified first for both groups. Afterwards, we analysed how stability and instability coexist within the network, exploring

https://doi.org/10.1017/S1366728925000185 Published online by Cambridge University Press

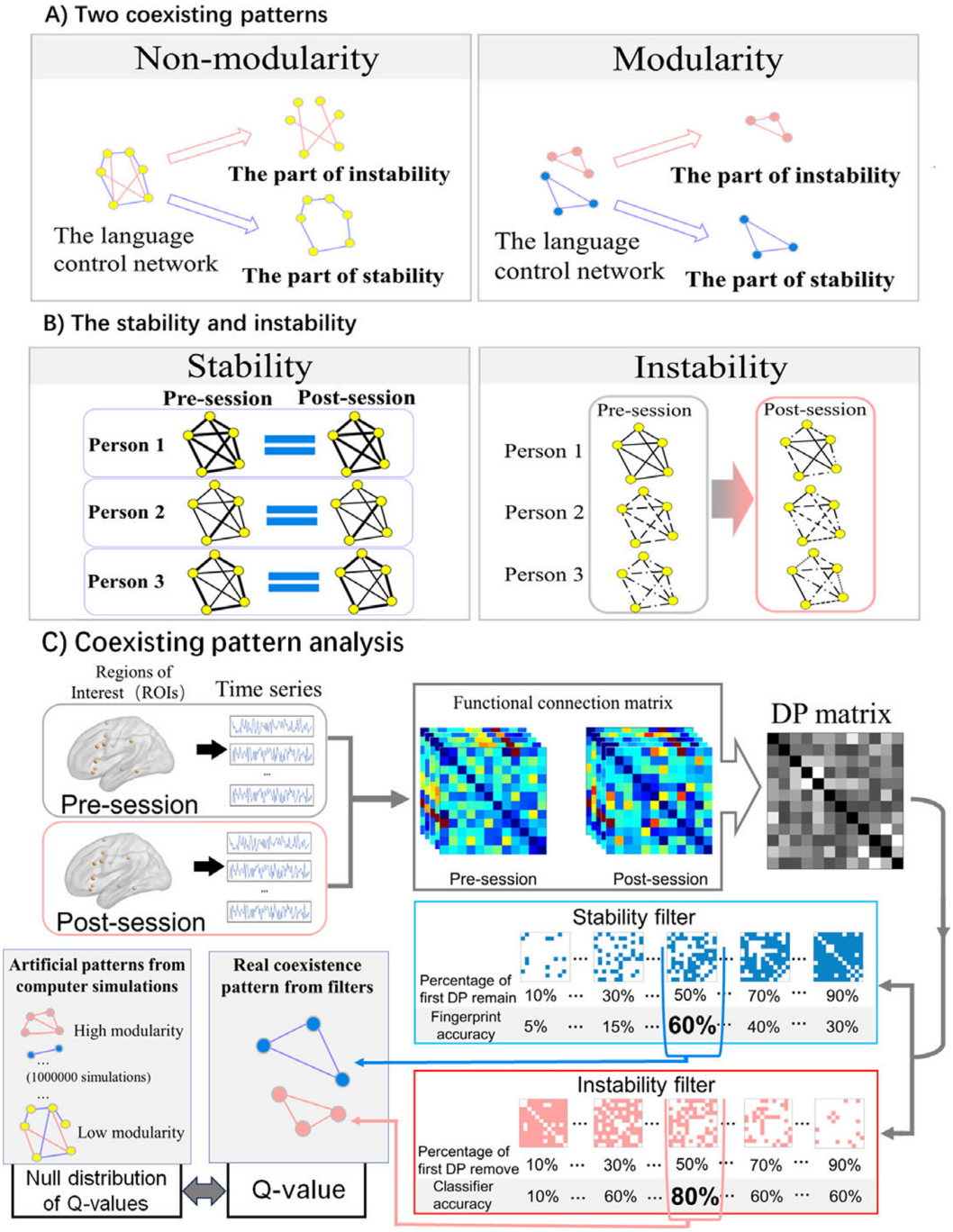


Figure 1. Schematic representation of the two key concepts of this study and the coexistence pattern analysis pipeline. (A) Schematic representation of the two possible coexistence patterns, with a non-modular schematic on the left and a modular schematic on the right. The dots represent different nodes (brain regions) and the connections between the dots represent edges (connections) between the nodes. (B) Schematic representation of stability and instability. The left side is stable, and the blue equal sign means that the left and right sides of the network are consistent. The right side is unstable, and the gradient-collared arrows indicate that there has been an overall change from the pre-session to the post-session. (C) A pipeline for the analysis of coexistence models. First, the BOLD signals are extracted for each region of interest within the network. The time series are then computed to obtain the functional connectivity matrix for the different sessions of each subject. The functional connectivity matrix is then employed to obtain the DP value matrix. Two filters are then used on the basis of the DP values. The upper blue filter is the stability filter, which is responsible for identifying the set of connections that best represents stability, and the lower blue filter is the instability filter, which is responsible for identifying the set of connections that best represents instability. The results of the filters are then aggregated to obtain the true coexistence patterns, and calculating the modularity coefficients (Q-values). Finally, the position of the Q-value of the true pattern in the null distribution constructed from the random pattern is calculated to measure the degree of modularity of the true pattern.

whether these properties manifest in modular or non-modular patterns. In addition, we performed the above analyses on the dorsal attention network, the default mode network and the whole-brain network for both groups, aiming to test whether the effect of L2 experience was specific to the language control network. We hypothesized that, due to the expansion of L2 classroom learning, the language control network in the English major group will demonstrate higher instability, lower stability and lower modularity

https://doi.org/10.1017/S1366728925000185 Published online by Cambridge University Press

in coexistence patterns. Conversely, the other three networks would exhibit no between-group differences due to the specificity of the effects of L2 experience.

2. Materials and methods

2.1. Participants

The English major group included 26 first-year undergraduate students majoring in English at South China Normal University (SCNU). Three participants were excluded because they did not participate in the post-test session, and another three were excluded because of poor data quality owing to excessive head movements. There were only 2 men among the remaining 20 participants, so to prevent the gender imbalance from affecting the results, we retained the data of 18 women (mean age = 19.61 years old [standard deviation (SD) = 1.16]) for the subsequent analyses. The English major group underwent two fMRI scans for approximately 1 year. During the year of L2 learning, the participants accessed naturalistic, varied and immersive L2 learning in real life, different from traditional L2 learning studies, in which specific training tasks (phonological/semantic choice tasks) under laboratory conditions were chosen (Qu et al., 2017; Wang et al., 2017). The average age at which they started learning English was 7.56 years old (SD = 2.95). This study was approved by the Research Ethics Committee of SCNU. All participants signed written informed consent forms before the experiment and received compensation for their time during the experiment.

Exactly 39 Chinese college students from the Southwest University Longitudinal Imaging Multimodal (SLIM) (Liu et al., 2017) were recruited as the non-English major group (mean age: 19.64 years old [SD: 0.97]). Similar to the English major group, only females were included in the non-English major group. They were all native Chinese speakers, and none were English majors. Compared with English majors, non-English majors received less classroom English learning in 1 year. They only accepted 1–2 h of weekly English classroom learning, which is far less in quantity and variety than the approximately 40 h of weekly bilingual learning for English majors.

There was no difference in age ($t = -0.099$, $p = 0.921$) and intelligence levels ($t = -0.46$, $p = 0.649$) between the English-majors and non-English majors. Data on the age of L2 acquisition for the non-English majors was missing because it was unavailable in the public database. However, the close rankings of the two comprehensive universities in the U.S. News & World Report Best Global Universities Rankings (Southwest University ranked 667, whereas SCNU ranked 770) and Shanghai Ranking's Best Chinese Universities Rankings (Southwest University ranked 72, whereas SCNU ranked 79) suggest that student L2 proficiency levels are relatively similar. Therefore, it is unlikely that the non-English major group at one university will have a higher proficiency level than the English major group at the other university. In addition, English majors are believed to have greater L2 proficiency than non-English majors.

2.2. Stimuli and experimental design

The English majors underwent two sessions of resting-state scanning within 1 year, once in the first semester of their first year (i.e. pre-session) and once in the third semester (i.e. post-session). The non-English major group also underwent two sessions of scanning within 1 year.

2.3. Imaging data acquisition

Resting-state fMRI scanning lasted for 8 min. During scanning, English majors were asked to close their eyes, rest, think of nothing and remain still. MRI images were acquired using a 3 T Siemens Trio scanner with a 12-channel phase-array head coil at the SCNU. Functional images were acquired using a T2*-weighted gradient-echo echo planar imaging (EPI) sequence with the following sequence parameters: repetition time (TR) = 2000 ms, time to echo (TE) = 30 ms, flip angle = 90°, field of view (FOV) = 204 × 204 mm^2, matrix = 64 × 64, slice thickness = 3.5 mm, interslice gap = 0.5 mm and voxel size = 3 × 3 × 3.5 mm^3. High-resolution brain structural images were acquired for all participants using a three-dimensional T1-weighted MP-RAGE sequence (TR = 1900 ms, TE = 2.52 ms, flip angle = 9°, FOV = 256 × 256 mm^2, matrix = 204 × 204, slice thickness = 1 mm and voxel size = 1 × 1 × 1 mm^3).

Images of the non-English major group were acquired using a 3 T Siemens Trio scanner with a 12-channel phase-array head coil at Southwest University. Functional images were obtained using a T2*-weighted gradient-echo EPI sequence (TR = 2000 ms, TE = 30 ms, flip angle = 90°, FOV = 220 × 220 mm^2, matrix = 64 × 64, slice thickness = 3.0 mm, interslice gap = 0.5 mm and voxel size = 3.4 × 3.4 × 3.0 mm^3).

2.4. Imaging data preprocessing

Data were preprocessed using the GRETNA toolbox based on SPM 12 (www.fil.ion.ucl.ac.uk/spm/software/spm12/) (Wang et al., 2015). The preprocessing steps included (1) removing the first 10 images, (2) slice timing correction, (3) realignment, (4) normalization to MNI space, (5) resampling to 3 mm isotropic voxels, (6) spatial smoothing with a 6 mm Full Width at Half Maximum Gaussian kernel, (7) removal of linear drift, (8) regressing out 24-parameter head motion profiles (Friston et al., 1996) and the global, white matter and cerebrospinal fluid signals and (9) temporal filtering with frequency of 0.008–0.083 Hz.

In the present study, different scanners with slightly varying parameters were used for the two groups. To ensure that the observed group differences are not artefacts of these variations, additional analyses were used. We used the method provided by Dietrich et al. (2007) to calculate the signal-to-noise ratio (SNR) of the two groups to measure the difference in image quality between the two groups. To more fully observe the SNR differences between the two groups, we used the poswer264 template to segment the whole brain. This template divides the whole brain into 264 regions, so we calculated SNRs for all 264 brain regions. After completing the preprocessing, independent samples *t*-tests were performed for each brain region. After Bonferroni correction for multiple comparisons, there were no significant between-group differences in the SNRs of all brain regions. We also used the Bayes factor (BF_{01}) to measure the SNR of the two groups. The alternative hypothesis was that there was no difference between the two groups. The mean BF_{01} value was 3.23, which suggests that substantial evidence exists to support that there is no difference in SNR between the two groups. In conclusion, based on the results of the *t*-test and Bayes factor, we concluded that there was no difference in image quality between the two groups.

2.5. Network construction

We selected 12 regions of interest (ROIs) to construct the language control network. The Left MFG (BA46), Pre-SMA, Left IFG(BA47),

https://doi.org/10.1017/S1366728925000185 Published online by Cambridge University Press

Right PrCG, Right Caudate, Left MTG, Left IFG (BA44), Right STG, Left MFG (BA9) and Left Caudate from a meta-analysis (Luk et al., 2012), dACC and Left IPL from previous studies (Barbeau et al., 2017; Green & Abutalebi, 2013; Liu et al., 2021a). dACC and Left IPL were selected based on two criteria: first, their presumed association with language control, and second, their documented involvement in long-term L2 training. All nodes were defined as spheres with a radius of 6 mm. The mask for the ROIs is based on the brain mask provided by GRETNA (Wang et al., 2015) toolbox. The coordinates and names of the brain regions in the language control network are presented in Table 1.

In addition, to contrast the results with those of the language control network, we used three other networks, the whole brain network, the default mode network and the dorsal attention network. The first two networks served as high-stability references (Finn et al., 2015; Jalbrzikowski et al., 2020; Kabbara et al., 2021). The dorsal attention network, which possesses a network size that is comparable to that of the language control network, can exert control over the effect of the network size. We chose the power264 template (Power et al., 2011) to segment the whole brain to 264 nodes. The default mode network contained 58 nodes and the dorsal attention network contained 11 nodes. The names and coordinates of each brain region contained in the default mode network and the dorsal attention network were defined by power264 template.

Functional connectivity matrices were analysed using the GRETNA (Wang et al., 2015) toolbox. The steps of functional connection network construction are as follows: (1) the BOLD time series of all voxels within each node was averaged; (2) the inter-nodal Pearson's correlation of BOLD time series was computed and (3) Fisher r-to-Z transformation.

2.6. Analysis of network stability and instability

The previous steps have completed the construction of the functional connectivity network; the next step is to measure the degree of stability and instability toward the functional network in general.

To comprehensively assess the characteristics of the language control network, we conducted two distinct analyses: stability assessment through functional connectome fingerprinting and evaluation of instability via a classifier approach.

Table 1. Node coordinates of language control network in MNI space

Node	Coordinate
Left middle temporal gyrus (L_MFG, BA46)	(−48, 49, 27)
Pre-supplementary motor area (Pre_SMA)	(2, 4, 64)
Left inferior frontal gyrus (L_IFG, BA47)	(−33, 25, −13)
Right precentral gyrus (R_PrCG)	(46, −5, −29)
Right Caudate	(17, 10, 11)
Left middle temporal gyrus (L_MTG)	(−54, −43, −11)
Left inferior frontal gyrus (L_IFG, BA44)	(−52, 21, 4)
Right superior temporal gyrus (R_STG)	(55, −19, −8)
Left middle frontal gyrus (L_MFG, BA9)	(−44, 8, 31)
Left Caudate	(−11, 22, −7)
Dorsal anterior cingulate cortex (dACC)	(0, 6, 44)
Left inferior parietal lobe (L_IPL, BA40)	(−54, −53, 36)

2.6.1. Functional connectome fingerprinting for stability assessment

We utilized functional connectome fingerprinting, a method that leverages the unique patterns of brain connectivity to identify individuals (Bari et al., 2019; Dufford et al., 2021; Finn et al., 2015; Ravindra et al., 2019). The more unique the connectome pattern is, and the less this unique pattern changes over time, the easier it is to recognize the connectome pattern of an individual participant, and the more accurate the connectome fingerprinting will be. Fingerprinting thus provides an index of the network's resistance to change over time, serving as a proxy for stability. A schematic of the connectome fingerprinting is shown on the left side of Figure 1B.

Fingerprinting was implemented through the following steps: (1) selecting one participant and computing the Pearson correlation to connectivity networks between this participant in the pre-session and all other participants in the post-session. If two networks from the same participant have the highest correlation coefficient, this participant's identification is successful; (2) calculating the accuracy of everyone and then adding them up to get a total accuracy (the accuracy of pre-post session); (3) computing another total accuracy after switching the pre- and post-session (the accuracy of post–pre session), and then computing the average value between the accuracy of pre–post session and the accuracy of post–pre session as the fingerprinting accuracy and (4) permutation testing: randomizing the index of the participants in pre-session and computing the total accuracy. This permutation test was repeated 1000 times. Statistical significance was defined as the percentage of the accuracy distribution higher than the original accuracy.

To further verify whether there is a difference in the fingerprinting accuracy between the English majors and the non-English major groups, we also tested the difference in their accuracy. First, the connections of the English major group and the non-English major group are mixed together and then reallocated back to the original matrix after being randomly disrupted. Then the connectome fingerprinting accuracy was calculated separately for the English majors and the non-English major groups after the permutation to get two accuracy rates, and finally the absolute value of the difference between these two accuracy rates was kept. A total of 1000 random permutations were performed, these 1000 differences form a random distribution. Finally, the rate at which the random distribution is higher than the original difference is calculated as the statistical significance.

The language control network, the dorsal attention network, the default mode network and the whole brain network all performed these calculations.

2.6.2. Network classifier for detecting instability

In the present study, we selected the classification accuracy of a Support Vector Machine (SVM) as an indicator of instability. This is because SVM fulfil the requirement of 'the opposite of stability'. Specifically, high stability, as measured using fingerprinting, implies significant interindividual variability, which remains stable over time. In contrast, instability implies small individual differences that vary significantly over time. The SVM can construct a hyperplane, separating the participants at different sessions in two places. The further the participants at different time points are from the hyperplane, the higher the classification accuracy. Therefore, when there is a difference between the pre- and post-sessions and the individual differences within the pre- and post-sessions are

https://doi.org/10.1017/S1366728925000185 Published online by Cambridge University Press

small, SVM has the best classification effect and the highest accuracy. A schematic of the SVM is shown on the right side of Figure 1B.

The specific process of using classifiers in the present study is as follows. The pre-session was labelled −1, whereas the post-session was labelled 1. Because of the small number of participants, leave-one-out cross-validation was used to evaluate the generalizability of the model (Tang et al., 2018). First, a participant's pre- or post-session data were employed as a test set, and then the model was trained using the remaining data. After obtaining the model, we tested whether it could correctly predict whether the test set belonged to the pre- or post-sessions. Analogously, each participant's pre- and post-session data were used as a test set to verify whether the model could successfully predict the pre- and post-sessions. The final success percentage was used as classification accuracy. Our classifier was based on the LIBSVM toolbox (Chang & Lin, 2011), with a linear kernel and otherwise default parameters. The permutation was repeated 1000 times to obtain a p-value for accuracy.

Additionally, the differences in classification accuracy between the English-majors and non-English-major groups were evaluated by permutation analysis. The permutation method and p-value calculations are consistent with Section 2.6.1.

The language control network, the dorsal attention network, the default mode network and the whole brain network all performed these calculations.

2.7. The coexisting pattern analyses

This study evaluated the stability using connectome fingerprinting and instability using classifiers. While the aforementioned analysis addresses the network as a whole and is unable to answer the question of how stability and instability coexist in a single network. It is as if knowing that 60% of someone's hair is black and 40% is white is not the same as knowing the pattern of co-existence of the two colours. Because it's possible that the back of the person's hair is completely black and the rest of the hair is white, or it is possible that the person has a full head of long black hair, but dyed the ends white, so each hair is also 60% black and 40% white. Similarly, the coexistence of stability and instability in a network cannot be determined by simply knowing the degree of stability and instability.

To study the coexistence patterns of stability and instability, it is first necessary to identify the two parts of the network that embody stability and instability. Our idea is to filter two connection sets from the whole network, the first one represents the stable part of the network and the other one represents the unstable part, and then determines whether the coexistence mode is modular or non-modular based on the morphology of the two sets.

Hence it is necessary to find the two connection sets that can represent stability and instability. Specifically, in the measure of the entire network earlier in this study, we used connectome fingerprinting to measure the stability of the entire network, so we also used connectome fingerprinting to find the set of connections that best represents stability. The connection set with the highest fingerprinting accuracy is considered the most stable connection set. Similarly, the connection set with the highest classification accuracy was considered to represent instability.

To filter faster, we employed the Differential Power (DP) algorithm. A set of connections containing a high proportion of connections with high DP values is more likely to exhibit stability, whereas a set of connections containing a large number of connections with low DP values is more likely to exhibit instability. Filtering using DP values is 25 times faster than exhaustive methods.

After filtering out the two connection sets, it is necessary to measure the degree of modularity of the coexistence patterns. In this study, to provide a more intuitive measure of the degree of modularity of coexisting patterns, computer simulations were used to generate a null distribution. In this way, the position of the real coexisting patterns in the null distribution can be used as an intuitive indicator of the degree of modularity of the coexisting patterns.

A series of analyses were performed to detect the coexistence pattern of stability and instability in the language control network. First, a DP matrix is calculated and then two filters are made based on the DP values. The filters find the connected set of stability and instability and finally the actual results obtained are compared with the simulated results to measure the modularity of the coexistence pattern (see Figure 1C). The language control network, the dorsal attention network, the default mode network and the whole-brain network were subjected to the above analyses.

2.7.1. The stability and instability filters

This step is comprised of three distinct components: the DP analysis and the two filters. The DP analysis is tasked with constructing the DP matrix, which serves as an index for the filters. The stability filter is responsible for identifying the set of connections that represent stability, while the instability filter is responsible for identifying the set of connections that represent instability.

2.7.1.1.Differential power analysis The DP analysis was implemented using custom code. We have largely maintained the algorithm of Finn et al. (2015) but adapted it slightly according to the data features in the present study to prevent infinite DP values. The steps for the DP analysis are as follows.

(1) Computing consistency of edges.

X^{R1} and X^{R2} indicate two normalized functional network matrices from pre- and post-session, respectively. i and j indicate two different subjects, e indicates edge and M is the total number of edges in one network. φ_{ii} indicates a vector of correspondence when the subject subscript is matched, whereas φ_{ij} and φ_{ji} indicate vectors of correspondence when subscript is unmatched. The specific formula is as follows:

$$\varphi_{ii}(e) = X_i^{R_1}(e) * X_i^{R_2}(e), e = 1,2,3,\cdots,M$$

$$\varphi_{ij}(e) = X_i^{R_1}(e) * X_j^{R_2}(e), e = 1,2,3,\cdots,M, i \neq j$$

$$\varphi_{ji}(e) = X_j^{R_1}(e) * X_i^{R_2}(e), e = 1,2,3,\cdots,M, i \neq j$$

(2) Computing empirical probability.

If $\varphi_{ii}(e)$ is equal to $\varphi_{ij}(e)$ or $\varphi_{ji}(e)$, edge e does not help to distinguish an individual from others. Therefore, an edge contributes to fingerprinting if it satisfies the following conditions only:

$\varphi_{ii}(e) > \varphi_{ij}(e)$, and $\varphi_{ii}(e) > \varphi_{ji}(e), i \neq j$

Based on the above conditions, we use an empirical probability $P_i(e)$ to qualify the DP:

$$P_i(e) = \frac{\left(\left|\varphi_{ii}(e) < \varphi_{ij}(e)\right| + \left|\varphi_{ii}(e) < \varphi_{ji}(e)\right|\right)}{2(n-1)}$$

where n denotes the number of participants. The lower the $P_i(e)$, the better the edge e distinguishes the subject i. Because the next

https://doi.org/10.1017/S1366728925000185 Published online by Cambridge University Press

step is logarithmic conversion, if $P_i(e)$ of a certain subject is zero, the outcome of the next step will be positive infinite. When n is relatively small, a positive infinity appears easily and covers the DP of the others. Extreme value makes subsequent analysis meaningless; therefore, we set a minimum of $P_i(e)[\min P_i(e)]$ to avoid positive infinity. To ensure $\min P_i(e)$ does not interfere with the analysis, its specific value must stratify the following conditions:

$$0 < \min P_i(e) < \frac{1}{2(n-1)}$$

If the two groups have different numbers of participants, the group with the lower number of participants will have a more generous range of $\min P_i(e)$, so the group with the higher number of participants should be used to determine the $\min P_i(e)$. Hence, the $\min P_i(e)$ should be lower than 0.13(1/39). We chose 0.01 as the $\min P_i(e)$.

(3) Logarithmic conversion.

The total DP of a certain edge across all subjects is defined by the DP measure as follows:

$$DP(e) = \sum_i \{-\ln(P_i(e))\}$$

Finally, a DP matrix of the edges was constructed.

2.7.1.2.The instability filter The instability filter was implemented using custom code based on the partial functions of SPM 12 and LIVSVM (Chang & Lin, 2011) toolbox. The steps for this filter are as follows.

(1) Constructing network masks with different sparsity.

The process involves creating masks with varying sparsity based on decreasing DP values for subsequent analysis. Starting with the DP matrix from the original network, edges are removed based on a percentage threshold (m%), with the remaining edges marked as 1 and removed ones as 0. This binary matrix represents a mask for a specific sparsity level. The vector m contains different sparsity, sorted from highest to lowest, including the screening process's start, end and step size.

For instance, an m value of 15 indicates that the top 15% of edges with the highest DP values are removed, retaining 85% of lower DP edges. Conversely, an m value of 70 removes the top 70%, leaving 30%. The Monte Carlo method is used to determine the probability of unsuccessful network formation due to edge removal. For the language control network, simulations showed a 99.999994% chance of network formation with 9% connections retained, increasing to 100% with 11%.

To preserve network information, a median value of 10% was chosen as the endpoint for m%, ensuring the final mask retains 10% of the lowest DP connections. Similar calculations led to the selection of 10% for the dorsal attention network and 1% for the default mode and whole brain networks.

The stable filter retains a consistent connection proportion. For the language control and dorsal attention networks, m ranges from 10 to 90. For the default mode and whole brain networks, m ranges from 1 to 99. Due to computational limitations, step sizes for m were adjusted: 1 for the language control and dorsal attention networks, and 0.1 for the default mode and whole brain networks.

Using these parameters, 81 masks were created for the language control and dorsal attention networks, and 981 masks for the default mode and whole brain networks, each representing a different level of sparsity.

(2) Classifier.

In the previous step, we completed the creation of masks for different DP sparsity, in this step, the masks are used to calculate the classification accuracy for different sparsity.

The initial step involves utilizing the initial value of the mask for m. This entails performing a dot product between all the functional connectivity matrices of each participant and the mask. This process yields a set of connections with a DP value below a specified threshold in the functional connectivity network. Subsequently, the functional connectivity matrix with a mask applied is computed in precisely the same manner as in Section 2.6.2, utilizing identical parameters and steps. This resulted in the classification accuracy corresponding to the first value of m. In the case of the language control network, this entailed the removal of connections in the initial 10% of DP values, with the remaining 90% retained, before calculating the classification accuracy for that specific set of connections. Similarly, the aforementioned calculations were performed for all values of m subsequently. This enabled the classification accuracy of the connection set with varying degrees of sparsity to be determined. Ultimately, the set of connections with the lowest number of connections and the highest classification accuracy is deemed to exemplify instability.

The aforementioned calculations were performed for the language control network, the dorsal attention network, the default mode network and the whole-brain network, according to the corresponding m value. All values of m were considered in these calculations.

2.7.1.3.The stability filter The stability filter was employed using custom code based on the partial functions of SPM 12 and previously described codes (Finn et al., 2015).

The steps for this filter are somewhat similar to those for the instability filter and are as follows:

(1) Constructing network masks of different sparsity.

The majority of the steps involved in the construction of masks are analogous to those employed in the instability filter, with the exception of the manner in which the initial m% of connections with DP values is treated. In particular, the instability filter necessitates a low DP value, thereby resulting in the removal of the top m % of connections with the highest DP value and the retention of the remainder. Conversely, the stability filter requires a high DP value, which consequently entails the retention of the top m% of connections and the removal of the remaining ones.

(2) Fingerprinting.

First, a specific mask is applied to the functional connectivity matrix, which is precisely the same step as the instability filter. Subsequently, the identical calculations outlined in Section 2.6.1 on the functional connectivity matrix with the mask applied must be performed to ascertain the fingerprint identification accuracy. Similarly, after applying masks corresponding to all values of m, the accuracy of fingerprint recognition is obtained for the set of connections corresponding to different sparsity. Ultimately, the connection set exhibiting the highest fingerprinting accuracy and comprising the fewest connections is deemed to represent the stability.

The aforementioned networks, namely the language control network, the dorsal attention network, the default mode network and the whole brain network, were all subjected to the aforementioned process.

2.7.2. Morphological coexistence pattern analysis

We conceptualize stable and unstable connection sets as two distinct communities within a network. In the network matrix, stable

https://doi.org/10.1017/S1366728925000185 Published online by Cambridge University Press

connections are marked '1', unstable ones '2' and those not belonging to either are '0'. This categorization divides the connections into two communities: Community 1 for stability and Community 2 for instability. The modularity algorithm then calculates a modularity coefficient (Q-value) to assess the separation of these communities (Newman, 2004).

To assess the modularity of the actual coexistence pattern, we generated a null distribution. This was done by randomly selecting connections from the network and assigning them to two communities, with the rest being coded as '0'. This random assignment simulates the distribution of connections across one million random networks. The Q-values from these networks form the null distribution.

In our study, the modularity of the actual coexistence pattern is determined by its Q-value's position within the null distribution. Modularity is quantified as the percentage of Q-values in the null distribution that is lower than the Q-value of the actual pattern. Modularity degree is expressed as a percentage for clarity and to enable comparison across networks of varying sizes. Percentages provide a more intuitive measure of modularity levels, allowing for a straightforward division into 'low', 'medium' and 'high' regions based on the null distribution. This contrasts with the Q-value, which lacks a consensus on what constitutes high or low modularity.

3. Results

3.1. Results for the test of stability and instability

In the language control network, fingerprinting accuracy was significantly higher in the non-English major group at 51.72% ($p < 0.01$) compared with the English major group at 36.11% ($p < 0.01$), with a significant difference between groups ($p < 0.01$). Classification accuracy was non-significant for both groups, with 50% for English majors ($p = 0.44$) and 27.59% for non-English majors ($p = 1$). But the difference between the two groups is marginally significant ($p = 0.052$).

For the dorsal attention network, fingerprinting accuracy was high and not significantly different between groups ($p = 0.24$), at 61.11% ($p < 0.01$) for English majors and 67.95% ($p < 0.01$) for non-English majors. Classification accuracy was also not significantly different from chance for both groups ($p = 0.37$), with 33.3% ($p = 1$) for English majors and 44.87% ($p = 0.76$) for non-English majors.

In the default mode network, fingerprinting accuracy was perfect for English majors at 100% ($p < 0.01$) and very high for non-English majors at 94.87% ($p < 0.01$), with no significant difference between groups ($p = 0.33$). Classification accuracy was non-significant for both groups ($p = 0.37$), with 11.11% ($p = 1$) for English majors and 29.48% ($p = 1$) for non-English majors.

In the whole-brain network, fingerprinting accuracy was 88.89% ($p < 0.01$) for English majors and 95.24% ($p < 0.01$) for non-English majors, with no significant difference ($p = 0.32$). Classification accuracy was 11.11% ($p = 1$) for English majors and 32.05% ($p = 0.99$) for non-English majors, with no significant difference between groups ($p = 0.38$).

In summary, significant differences in accuracy were observed only in the language control network between English majors and non-major groups, with no such differences in other networks. These results are depicted in Figure 2A.

3.2. Results of the morphological coexistence pattern analysis

The first is the filtering results for the language control network. For the English majors, 30% of connections with the lower DP have the highest classification accuracy (58%), and 58% of connections with higher DP has the highest fingerprinting accuracy (64%) (see in Figure 2B). For the non-English majors, 10% of connections with lower DP have the highest classification accuracy (62%) and 58% of connections with higher DP has the highest fingerprinting accuracy (57%) (see in Figure 2C).

The analysis of coexistence patterns within the language control network indicates that both the stability and instability connection sets for English majors encompass all brain regions, whereas only the stability set for non-English majors does so. The English majors' coexistence pattern modularity coefficient ($Q = 0.1592$) surpasses 26.84% of the computer-simulated random patterns, whereas the non-English majors' ($Q = 0.174$) exceeds 67.44%, with a group difference of 40.6%. According to the criteria of 'high, medium and low' modularity, the English majors fall into the 'low modularity area', while non-English majors are in the 'high modularity area'. The modularity coefficients for the null distribution of both groups are depicted in Figure 2D.

In the dorsal attention network, the English majors' coexistence pattern modularity ($Q = 0.1346$) is >4.5% of the random patterns, and the non-English majors' ($Q = 0.13457$) is 4.41%, with a minute group difference of 0.09%, both classified as low modularity. For the default mode network, the English majors' pattern ($Q = 0.2254$) exceeds the random patterns at 98.2%, and the non-English majors' ($Q = 0.2251$) at 98.0%, with a negligible difference of 0.2%, both in the high modularity region. Finally, the whole brain network's coexistence pattern modularity for English majors ($Q = 0.237$) exceeds the random patterns by 96.72%, and for non-English majors ($Q = 0.235$) by 93.35%, with a group difference of 3.37%, both in the high modularity region. The degree of modularity for the coexistence patterns of each network is detailed in Figure 2E.

In conclusion, the language control network shows a considerable difference in coexistence patterns between English and non-English majors, with English majors displaying low modularity and non-English majors' high modularity. The dorsal attention network is characterized by low modularity, while the default mode network and the whole brain network exhibit high modularity, with small between-group differences.

4. Discussion

The present study employed rs-fMRI to delve into the stability and instability of the language control network, with a particular focus on comparing the differences between Chinese university students majoring in English and those not. Our findings indicate that the language control network of the English-major group demonstrated lower stability and higher instability compared with the non-English major group. In addition, the English major group showed less modularity than the non-English major group in terms of coexisting patterns of stability and instability in the language control network. These two distinct differences were not observed in the other three networks, suggesting that L2 learning experience may extensively influence the function of the language control network. These results not only offer new insights into the neural mechanisms of language control in the bilingual brain but also highlight the potential role of L2 learning in shaping the language control network.

https://doi.org/10.1017/S1366728925000185 Published online by Cambridge University Press

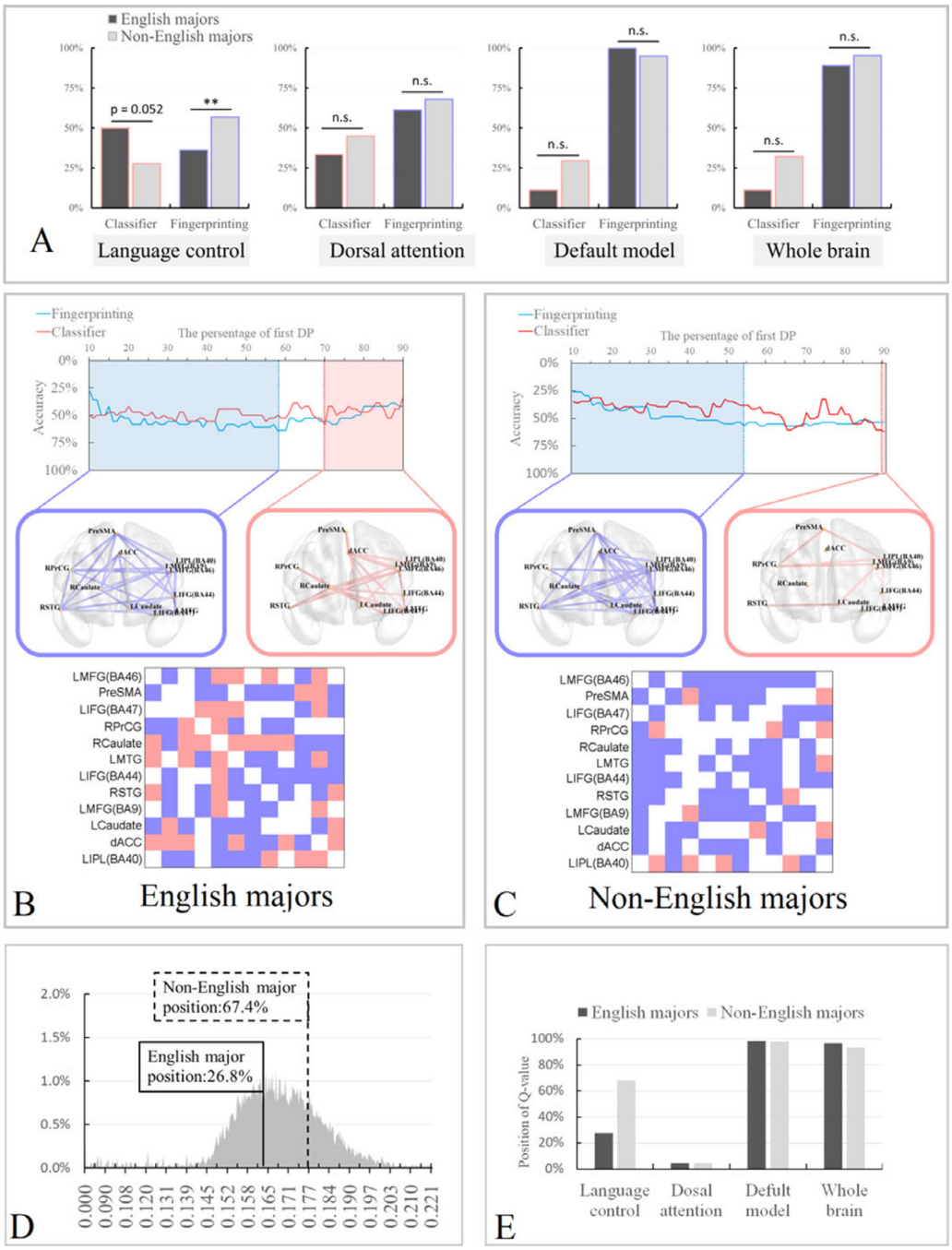


Figure 2. The results of the calculation of the stability and instability of each network; the filtering process of the language control network and the specific filtering results; the degree of modularity of the coexistence patterns of each network. (A) The results of the stability and instability tests for the four networks are presented herewith. **p < 0.01; n.s. indicates a non-significant result. (B) The results of the filtering process applied to the language control network for the group of English majors. The red and blue lines in the line graph represent the accuracy rates associated with connection sets of varying sparsity. The connection set corresponding to the red region exhibits the highest classification accuracy, while the connection set corresponding to the blue region demonstrates the highest fingerprinting accuracy. The connections corresponding to the red and blue regions are illustrated in the brain below each one. The matrix at the bottom illustrates the coexistence of stability and instability, with blue indicating stability and red indicating instability. The brain images were generated using the BrainNet Viewer software (Xia et al., 2013). (C) Filtering results of the language control network for the English major group. (D) The degree of modularity in the coexistence patterns of language control network in the English majors and non-English major groups. The grey area represents the null distribution, which comprises the modularity coefficients of the coexistence patterns that have been randomly simulated by the computer. The solid lines demarcate the position of the English major group, while the dashed lines demarcate the position of the non-English major group. (E) The degree of modularity of the coexistence patterns of the individual networks in the English and non-English groups are illustrated. This study measures the degree of modularity using the position of the modularity coefficients in their respective null distributions.

4.1. Stability and instability in the language control network

This section delves into three critical dimensions of the language control network: stability, instability and a comparative analysis of these aspects between English majors and non-English major groups.

4.1.1. Stability in the language control network

The study's findings on the stability of the whole brain in healthy adults align with prior research, confirming the high stability of the resting-state functional network (Finn et al., 2015; Liu et al., 2018; Ravindra et al., 2019). The whole brain fingerprinting

https://doi.org/10.1017/S1366728925000185 Published online by Cambridge University Press

accuracy for the two groups was 88.89% and 95.24%, fitting within the 80%–95% range reported previously. The default mode network's stability is comparable to the whole brain, highlighting its individualistic and stable nature. Notably, the language control network's stability, at 36.11% and 51.72%, is significantly lower, suggesting it is less stable than the whole-brain and default mode networks. However, it is not as low as the infant whole brain's stability, which is only 26.6% (Dufford et al., 2021), indicating a moderate stability for the language control network. The number of nodes in the dorsal attention network is similar to that of the language control network. However, the fingerprinting accuracy in the former is 1.47 times higher than that in the latter. This suggests that the language control network's stability is not solely node-dependent. Overall, the language control network exhibits a moderate degree of stability.

4.1.2. Instability of the language control network

While the classifiers in this study did not yield significant results, this does not imply a lack of meaningful insights. It is crucial to clarify that the random levels for classification accuracy and fingerprint recognition accuracy are distinct measures. The random level for fingerprint recognition, at 5.56% (1/18) and 2.56% (1/39), is inversely related to the number of participants. In contrast, the random level for classification accuracy is 50%, based on the number of scans to be classified.

The leave-one-out cross-validation method, a staple in machine learning for small datasets, plays a pivotal role here. This method isolates a single participant's pre- or post-session for testing while training on the remaining data. It is particularly sensitive to individual differences. If there are substantial individual variations, the isolated test data may significantly deviate from the trained model, potentially leading to classification accuracies that fall below the random level. This scenario is exacerbated when the differences between pre- and post-sessions are minimal, increasing the likelihood of individual classification failures.

The classification accuracy obtained through leave-one-out cross-validation is thus an inverse indicator of fingerprint recognition accuracy. The greater the individual differences and the smaller the test differences, the higher the fingerprint recognition rate and the lower the classification accuracy. The antagonistic relationship between these two metrics was confirmed with a Spearman correlation coefficient of -0.76 ($p = 0.03$), underscoring the antagonistic relationship between stability and instability.

In essence, the non-significant classification accuracies in this study can be attributed to the algorithmic properties of leave-one-out cross-validation and its sensitivity to individual differences. Despite this, the higher classification accuracy of the language control network in the English major group, compared with other networks, suggests a greater degree of instability or change following a year of bilingual training.

4.1.3. Differences between English major and non-English major groups.

The study identified significant differences in the language control network between English major and non-English major, with fingerprinting differences being statistically significant and classification accuracy showing marginal significance. The other three networks did not exhibit notable intergroup differences. Our analysis ruled out image quality as a factor, as there were no significant differences in the SNR between groups. Given the similarity in age, intelligence and education level among the female participants, the differences in the language control network are likely not due to these variables. The distinct findings for the language control network, in contrast to the whole-brain network, suggest that the observed differences are specific to language control and not a reflection of broader brain network differences. The lack of differences in the default mode and dorsal attention networks indicates that the observed differences in the language control network are not driven by resting-state monitoring or network size.

In conclusion, the differences in the language control network are likely due to the specific demands of language control training. English majors, having received more extensive L2 training, showed adaptive changes in their language control networks, leading to increased instability and decreased stability. This suggests that the greater language control exercised by English majors results in more pronounced changes in brain regions and connections, a finding supported by existing research (Abutalebi & Rietbergen, 2014; Barbeau et al., 2017; Bialystok, 2014, 2021; Kroll & Chiarello, 2016; Li et al., 2014; Liu et al., 2021a; Pliatsikas, 2020; Tu et al., 2015; Zou et al., 2012).

In the two scans of this study, English majors studied English for 40 hours per week, including speaking, writing, dialogue, reading, speech, translation, debate and the study of historical and literary masterpieces of British and American countries. According to the curriculum of the school to which the group of non-English majors belongs, most of the non-English majors receive only 1 or 2 h of public classes per week, and the public English classes available to the students of this school are not comprehensive, but include only a part of English training, such as only vocabulary, writing or reading. In conclusion, since the data for the non-English major group in this study came from a public database, and information on their age of L2 acquisition, daily use and L2 proficiency was not collected, it is not possible to directly ensure that there is a significant difference between the English and non-English major groups in terms of L2 experience and proficiency, but based on the content of their core curriculum, it is assumed that accepting the existence of a difference between the two groups may be more reasonable than denying the existence of a difference between the two groups.

Overall, there was a moderate degree of stability and instability in the language control network, with greater instability for English majors and greater stability for non-English majors, a difference that may be due to the longer and more varied L2 training received by the English major group.

4.2. Morphological coexistence pattern

Stability and instability have been discussed in the previous section, but the metrics corresponding to the two properties do not reflect how the two properties coexist in a network, that is how a network can be both stable and unstable at the same time.

The results show that both the whole-brain network and the dorsal attention network exhibit a high degree of modularity in their patterns of coexistence, and that the differences between the English majors and non-English major groups are small. This suggests that the connections occupied by stability and instability in these two networks constitute two sharply delimited modules. In other words, stability and instability each occupy two parts of the brain, and these two parts overlap less. This could mean that only a very small number of brain regions changed in both networks, with the majority remaining stable. The pattern for the dorsal attentional network was quite the opposite, with the dorsal attention network showing very low modularity in its coexistence pattern, and much

https://doi.org/10.1017/S1366728925000185 Published online by Cambridge University Press

lower than the language control network. This implies that the dorsal attention network occupied a large number of brain regions for each of stability and instability. This may indicate that the dorsal attention network employs not a pattern in which some brain regions remain stable while others change, but rather a pattern of activity in which some patterns of connectivity throughout the network are responsible for change while others remain stable. The language control network, on the other hand, differed considerably between the English majors and non-English major groups: the non-English major group was more modular and the English major group was less modular. This apparent difference between groups was not observed on the other three networks. This implies that the pattern of coexistence of stability and instability in language control changes with L2 experience. The coexistence pattern in the non-English major group tended to show stability in most brain regions and instability in a few, whereas the English majors tended more to show stability in one connectivity pattern and instability in the other. We suggest that this may be suggestive of the process by which L2 experience influences language control networks, especially functional networks.

The L2 learning of non-English majors is less in quantity and quality. This is similar to the initial exposure stage in the DRM. The DRM suggested that a specific part of the node was first affected by L2 experience during the initial exposure stage. The present study found that instability in non-English majors did not involve all language control network nodes. In contrast, instability in English majors involved all nodes rather than part of the nodes. This may indicate that the influence of L2 experience is not limited to certain brain regions but to all nodes of the language control network. Synthesizing the coexistence patterns in the language control networks of English and non-English majors, we propose a hypothesis regarding the impact of classroom L2 learning on the language control network. The mechanism through which L2 learning affects the language control network may be similar to the gradual construction of multiple expressways. The brain regions are similar to many cities. For non-English majors, the year they received less L2 learning may be similar to the initial stages of building a highway network. During the initial stages of expressway construction, only a few cities are connected by expressways. For English majors, the situation is similar to the later stages of an expressway network construction project. In the later stages of expressway construction, specific highways are connected to all cities. However, not every city is directly connected; some cities are only indirectly connected to other cities because of construction costs or other reasons. Initially, only a few connections in certain brain regions were affected by L2 learning. Eventually, some connections in each brain region were affected. Therefore, the stable and unstable parts shared some nodes in non-English majors, whereas they shared all nodes in English majors. If we think of the nodes as chemical elements and the connections as bonds, the stable and unstable parts are similar to a pair of isomers in chemistry (compounds with identical chemical formulae but different structures) (Regalado et al., 2013).

This result is inconsistent with that of the DRM. In the DRM, only a few specific brain regions change at each stage. This disparity may be due to the complex mapping of changes in brain function and structure. The DRM focused on the dynamic restructuring of gray and white matter, whereas we focused on the functional connectome. Complex mapping between brain structure and functional connectivity may be responsible for this inconsistency.

4.3. The temporal complexity of the language control network

Tononi et al. (1994) proposed that a brain network is complex when neither segregated nor integrated. The more balanced the two features are, the higher the complexity of the network as a whole (Bassett & Gazzaniga, 2011; Marshall et al., 2016). Integration and separation, as well as stability and instability, are at the two ends of a continuum. Therefore, we can preliminarily propose 'temporal complexity' using analogy with morphological or spatial complexity. Specifically, the brain network is a temporal complex that is neither stable nor unstable. In the present study, the language control network of English major may be considered to have high temporal complexity. However, the present study is only a preliminary exploration and does not provide very strong evidence. The temporal complexity of brain remains to be further explored.

4.4. Limitations

First, the present study had a small sample size. This may have weakened the generalizability and robustness of our results and conclusions. Second, learning for over 1 year was affected by many irrelevant variables. Randomization and other methods were used to eliminate interference; however, there is no guarantee that classroom L2 learning was the only independent factor. A methodological limitation of this study is the lack of direct measurement and control over participants' second-language experience and proficiency. This omission may have introduced additional variability, complicating the interpretation of true sources and mechanisms behind group differences in stability and instability. We acknowledge this and aim to address it in future studies by incorporating direct assessments of second-language experience and proficiency or employing more sophisticated control methods.

5. Conclusions

This study shows the coexistence of stability and instability in the language control network. Specifically, the language control network exhibited a moderate level of both stability and instability, with these properties coexisting in a non-modular fashion. Notably, English majors exhibited a more significant non-modularity compared with non-English majors.

Data availability statement. The dataset analysed in the present study are available from the corresponding authors through the Harvard Dataverse (https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ONRHXA).

Acknowledgments. The authors would like to thank Editage for the technical editing of the manuscript.

Funding statement. This study was funded by the National Natural Sciences Foundation of China (32371114, 32471100), Guangdong Basic and Applied Basic Research Foundation (2024A1515012354), Research Center for Brain Cognition and Human Development, Guangdong, China (2024B0303390003) and Striving for the First-Class, Improving Weak Links and Highlighting Features (SIH) Key Discipline for Psychology in South China Normal University.

Competing interest. The authors have no competing interests to declare.

References

Abutalebi, J., Della Rosa, P. A., Ding, G., Weekes, B., Costa, A., & Green, D. W. (2013). Language proficiency modulates the engagement of cognitive

control areas in multilinguals. *Cortex*, **49**(3), 905–911. https://doi.org/10.1016/j.cortex.2012.08.018

Abutalebi, J., & **Green, D. W**. (2016). Neuroimaging of language control in bilinguals: Neural adaptation and reserve. *Bilingualism: Language and Cognition*, **19**(4), 689–698. https://doi.org/10.1017/S1366728916000225

Abutalebi, J., & **Rietbergen, M. J**. (2014). Neuroplasticity of the bilingual brain: Cognitive control and reserve. *Applied Psycholinguistics*, **35**(5), 895–899. https://doi.org/10.1017/S0142716414000186

Alotaibi, S., **Alsaleh, A.**, **Wuerger, S.**, & **Meyer, G**. (2023). Rapid neural changes during novel speech-sound learning: An fMRI and DTI study. *Brain and Language*, **245**, 105324. https://doi.org/10.1016/j.bandl.2023.105324

Amico, E., & **Goñi, J**. (2018). The quest for identifiability in human functional connectomes. *Scientific Reports*, **8**(1), 8254. https://doi.org/10.1038/s41598-018-25089-1

Antoniou, M. (2019). the advantages of bilingualism debate. *Annual Review of Linguistics*, **5**(1), 395–415. https://doi.org/10.1146/annurev-linguistics-011718-011820

Barbeau, E. B., **Chai, X. J.**, **Chen, J.-K.**, **Soles, J.**, **Berken, J.**, **Baum, S.**, **Watkins, K. E.**, & **Klein, D**. (2017). The role of the left inferior parietal lobule in second language learning: An intensive language training fMRI study. *Neuropsychologia*, **98**, 169–176. https://doi.org/10.1016/j.neuropsychologia.2016.10.003

Bari, S., **Amico, E.**, **Vike, N.**, **Talavage, T. M.**, & **Goñi, J**. (2019). Uncovering multi-site identifiability based on resting-state functional connectomes. *NeuroImage*, **202**, 115967. https://doi.org/10.1016/j.neuroimage.2019.06.045

Bassett, D. S., & **Gazzaniga, M. S**. (2011). Understanding complexity in the human brain. *Trends in Cognitive Sciences*, **15**(5), 200–209. https://doi.org/10.1016/j.tics.2011.03.006

Baum, S., & **Titone, D**. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Applied Psycholinguistics*, **35**(5), 857–894. https://doi.org/10.1017/S0142716414000174

Bialystok, E. (2014). Neuroplasticity as a model for bilingualism: Commentary on Baum and Titone. *Applied Psycholinguistics*, **35**(5), 899–902. https://doi.org/10.1017/S0142716414000198

Bialystok, E. (2021). Bilingualism: Pathway to cognitive reserve. *Trends in Cognitive Sciences*, **25**(5), 355–364. https://doi.org/10.1016/j.tics.2021.02.003

Branzi, F. M., **Calabria, M.**, **Boscarino, M. L.**, & **Costa, A**. (2016). On the overlap between bilingual language control and domain-general executive control. *Acta Psychologica*, **166**, 21–30. https://doi.org/10.1016/j.actpsy.2016.03.001

Caffarra, S., **Molinaro, N.**, **Davidson, D.**, & **Carreiras, M**. (2015). Second language syntactic processing revealed through event-related potentials: An empirical review. *Neuroscience & Biobehavioral Reviews*, **51**, 31–47. https://doi.org/10.1016/j.neubiorev.2015.01.010

Cao, B., **Guo, Y.**, **Lu, M.**, **Wu, X.**, **Deng, F.**, **Wang, J.**, & **Huang, R**. (2024). The long-term intensive gymnastic training influences functional stability and integration: A resting-state fMRI study. *Psychology of Sport and Exercise*, **74**, 102678. https://doi.org/10.1016/j.psychsport.2024.102678

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

Chang, C., & **Lin, C**. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, **2**(3), 1–27. https://doi.org/10.1145/1961189.1961199

Chen, B., **Xu, T.**, **Zhou, C.**, **Wang, L.**, **Yang, N.**, **Wang, Z.**, **Dong, H.-M.**, **Yang, Z.**, **Zang, Y.-F.**, **Zuo, X.-N.**, & **Weng, X.-C**. (2015). Individual variability and test-retest reliability revealed by ten repeated resting-state brain scans over one month. *PLOS ONE*, **10**(12), e0144963. https://doi.org/10.1371/journal.pone.0144963

Declerck, M., **Koch, I.**, & **Philipp, A. M**. (2015). The minimum requirements of language control: Evidence from sequential predictability effects in language switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **41**(2), 377–394. https://doi.org/10.1037/xlm0000021

DeLuca, V., **Voits, T.**, **Ni, J.**, **Carter, F.**, **Rahman, F.**, **Mazaheri, A.**, **Krott, A.**, & **Segaert, K**. (2024). Mapping individual aspects of bilingual experience to adaptations in brain structure. *Cerebral Cortex*, **34**(2), bhae029. https://doi.org/10.1093/cercor/bhae029

Dietrich, O., **Raya, J. G.**, **Reeder, S. B.**, **Reiser, M. F.**, & **Schoenberg, S. O**. (2007). Measurement of signal-to-noise ratios in MR images: Influence of multichannel coils, parallel imaging, and reconstruction filters. *Journal of Magnetic Resonance Imaging*, **26**(2), 375–385. https://doi.org/10.1002/jmri.20969

Dufford, A. J., **Noble, S.**, **Gao, S.**, & **Scheinost, D**. (2021). The instability of functional connectomes across the first year of life. *Developmental Cognitive Neuroscience*, **51**, 101007. https://doi.org/10.1016/j.dcn.2021.101007

Fedorenko, E., & **Thompson-Schill, S. L**. (2014). Reworking the language network. *Trends in Cognitive Sciences*, **18**(3), 120–126. https://doi.org/10.1016/j.tics.2013.12.006

Finn, E. S., **Shen, X.**, **Scheinost, D.**, **Rosenberg, M. D.**, **Huang, J.**, **Chun, M. M.**, **Papademetris, X.**, & **Constable, R. T**. (2015). Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity. *Nature Neuroscience*, **18**(11), 1664–1671. https://doi.org/10.1038/nn.4135

Friston, K. J., **Williams, S.**, **Howard, R.**, **Frackowiak, R. S. J.**, & **Turner, R.** (1996). Movement-Related effects in fMRI time-series. *Magnetic Resonance in Medicine*, **35**(3), 346–355. https://doi.org/10.1002/mrm.1910350312

Green, D. W., & **Abutalebi, J**. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, **25**(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Jalbrzikowski, M., **Liu, F.**, **Foran, W.**, **Klei, L.**, **Calabro, F. J.**, **Roeder, K.**, **Devlin, B.**, & **Luna, B**. (2020). Functional connectome fingerprinting accuracy in youths and adults is similar when examined on the same day and 1.5-years apart. *Human Brain Mapping*, **41**(15), 4187–4199. https://doi.org/10.1002/hbm.25118

Jiao, L., **Meng, N.**, **Wang, Z.**, **Schwieter, J. W.**, & **Liu, C**. (2022). Partially shared neural mechanisms of language control and executive control in bilinguals: Meta-analytic comparisons of language and task switching studies. *Neuropsychologia*, **172**, 108273. https://doi.org/10.1016/j.neuropsychologia.2022.108273

Jiao, L., **Schwieter, J. W.**, & **Liu, C**. (2025a). Comprehension-based language switching in experienced and newly learned languages: Evidence from induced brain oscillations. *Journal of Neurolinguistics*, **73**, 101232. https://doi.org/10.1016/j.jneuroling.2024.101232

Jiao, L., **Wang, Z.**, **Duan, X.**, **Yu, Y.**, & **Liu, C.** (2025b). Linking context to language switching: Effects of background noise on bilingual language comprehension. *Behavioral Sciences*, **15**(1), 60. https://doi.org/10.3390/bs15010060

Kabbara, A., **Paban, V.**, & **Hassan, M**. (2021). The dynamic modular fingerprints of the human brain at rest. *NeuroImage*, **227**, 117674. https://doi.org/10.1016/j.neuroimage.2020.117674

Kang, C., **Fu, Y.**, **Wu, J.**, **Ma, F.**, **Lu, C.**, & **Guo, T**. (2017). Short-term language switching training tunes the neural correlates of cognitive control in bilingual language production: Language training modulates cognitive control. *Human Brain Mapping*, **38**(12), 5859–5870. https://doi.org/10.1002/hbm.23765

Kaufmann, T., **Alnæs, D.**, **Doan, N. T.**, **Brandt, C. L.**, **Andreassen, O. A.**, & **Westlye, L. T**. (2017). Delayed stabilization and individualization in connectome development are related to psychiatric disorders. *Nature Neuroscience*, **20**(4), 513–515. https://doi.org/10.1038/nn.4511

Korenar, M., **Treffers-Daller, J.**, & **Pliatsikas, C**. (2023). Dynamic effects of bilingualism on brain structure map onto general principles of experience-based neuroplasticity. *Scientific Reports*, **13**, 3428. https://doi.org/10.1038/s41598-023-30326-3

Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and Language*, **109**(2–3), 68–74. https://doi.org/10.1016/j.bandl.2008.06.002

Kroll, J. F., & **Chiarello, C**. (2016). Language experience and the brain: Variability, neuroplasticity, and bilingualism. *Language, Cognition and Neuroscience*, **31**(3), 345–348. https://doi.org/10.1080/23273798.2015.1086009

Li, P., **Legault, J.**, & **Litcofsky, K. A**. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex*, **58**, 301–324. https://doi.org/10.1016/j.cortex.2014.05.001

Liu, C., **de Bruin, A.**, **Jiao, L.**, **Li, Z.**, & **Wang, R**. (2021a). Second language learning tunes the language control network: A longitudinal fMRI study. *Language, Cognition and Neuroscience*, **36**(4), 462–473. https://doi.org/10.1080/23273798.2020.1856898

Liu, C., **Jiao, L.**, **Li, Z.**, **Timmer, K.**, & **Wang, R**. (2021b). Language control network adapts to second language learning: A longitudinal rs-fMRI study. *Neuropsychologia*, **150**, 107688. https://doi.org/10.1016/j.neuropsychologia.2020.107688

https://doi.org/10.1017/S1366728925000185 Published online by Cambridge University Press

Liu, C., **Jiao, L.**, **Timmer, K.**, & **Wang, R**. (2021c). Structural brain changes with second language learning: A longitudinal voxel-based morphometry study. *Brain and Language*, **222**, 105015. https://doi.org/10.1016/j.bandl.2021.105015

Liu, J., **Liao, X.**, **Xia, M.**, & **He, Y**. (2018). Chronnectome fingerprinting: Identifying individuals and predicting higher cognitive functions using dynamic brain connectivity patterns. *Human Brain Mapping*, **39**(2), 902–915. https://doi.org/10.1002/hbm.23890

Liu, L., **Yan, X.**, **Li, H.**, **Gao, D.**, & **Ding, G**. (2020). Identifying a supramodal language network in human brain with individual fingerprint. *NeuroImage*, **220**, 117131. https://doi.org/10.1016/j.neuroimage.2020.117131

Liu, W., **Wei, D.**, **Chen, Q.**, **Yang, W.**, **Meng, J.**, **Wu, G.**, **Bi, T.**, **Zhang, Q.**, **Zuo, X.-N.**, & **Qiu, J**. (2017). Longitudinal test-retest neuroimaging data from healthy young adults in southwest China. *Scientific Data*, **4**(1), 170017. https://doi.org/10.1038/sdata.2017.17

Luk, G., **Green, D. W.**, **Abutalebi, J.**, & **Grady, C**. (2012). Cognitive control for language switching in bilinguals: A quantitative meta-analysis of functional neuroimaging studies. *Language and Cognitive Processes*, **27**(10), 1479–1488. https://doi.org/10.1080/01690965.2011.613209

Marshall, N., **Timme, N. M.**, **Bennett, N.**, **Ripp, M.**, **Lautzenhiser, E.**, & **Beggs, J. M.** (2016). Analysis of power laws, shape collapses, and neural complexity: New techniques and MATLAB support via the NCC toolbox. *Frontiers in Physiology*, **7**. https://doi.org/10.3389/fphys.2016.00250

Munsell, B. C., **Gleichgerrcht, E.**, **Hofesmann, E.**, **Delgaizo, J.**, **McDonald, C. R.**, **Marebwa, B.**, **Styner, M. A.**, **Fridriksson, J.**, **Rorden, C.**, **Focke, N. K.**, **Gilmore, J. H.**, & **Bonilha, L**. (2020). Personalized connectome fingerprints: Their importance in cognition from childhood to adult years. *NeuroImage*, **221**, 117122. https://doi.org/10.1016/j.neuroimage.2020.117122

Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E*, **69**(6), 066133. https://doi.org/10.1103/PhysRevE.69.066133

Pliatsikas, C. (2020). Understanding structural plasticity in the bilingual brain: The dynamic restructuring model. *Bilingualism: Language and Cognition*, **23**(2), 459–471. https://doi.org/10.1017/S1366728919000130

Połczyńska, M. M., & **Bookheimer, S. Y**. (2021). General principles governing the amount of neuroanatomical overlap between languages in bilinguals. *Neuroscience & Biobehavioral Reviews*, **130**, 1–14. https://doi.org/10.1016/j.neubiorev.2021.08.005

Power, J. D., **Cohen, A. L.**, **Nelson, S. M.**, **Wig, G. S.**, **Barnes, K. A.**, **Church, J. A.**, **Vogel, A. C.**, **Laumann, T. O.**, **Miezin, F. M.**, **Schlaggar, B. L.**, & **Petersen, S. E**. (2011). Functional network organization of the human brain. *Neuron*, **72**(4), 665–678. https://doi.org/10.1016/j.neuron.2011.09.006

Qu, J., **Qian, L.**, **Chen, C.**, **Xue, G.**, **Li, H.**, **Xie, P.**, & **Mei, L**. (2017). Neural pattern similarity in the left IFG and fusiform is associated with novel word learning. *Frontiers in Human Neuroscience*, **11**, 424. https://doi.org/10.3389/fnhum.2017.00424

Ravindra, V., **Drineas, P.**, & **Grama, A**. (2019). Constructing compact brain connectomes for individual fingerprinting. *arXiv:1805.08649 [Cs]*. http://arxiv.org/abs/1805.08649

Regalado, E. L., **Schafer, W.**, **McClain, R.**, & **Welch, C. J**. (2013). Chromatographic resolution of closely related species: Separation of warfarin and hydroxylated isomers. *Journal of Chromatography A*, **1314**, 266–275. https://doi.org/10.1016/j.chroma.2013.07.092

Sander, K., **Chai, X.**, **Barbeau, E. B.**, **Kousaie, S.**, **Petrides, M.**, **Baum, S.**, & **Klein, D**. (2023). Interhemispheric functional brain connectivity predicts new language learning success in adults. *Cerebral Cortex*, **l33**(4), 1217–1229. https://doi.org/10.1093/cercor/bhac131

Spear, L. P. (2013). Adolescent neurodevelopment. *Journal of Adolescent Health*, **52**(2), S7–S13. https://doi.org/10.1016/j.jadohealth.2012.05.006

Tang, H., **Lu, X.**, **Cui, Z.**, **Feng, C.**, **Lin, Q.**, **Cui, X.**, **Su, S.**, & **Liu, C**. (2018). Resting-state functional connectivity and deception: Exploring individualized deceptive propensity by machine learning. *Neuroscience*, **395**, 101–112. https://doi.org/10.1016/j.neuroscience.2018.10.036

Tononi, G., **Sporns, O.**, & **Edelman, G. M**. (1994). A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences*, **91**(11), 5033–5037. https://doi.org/10.1073/pnas.91.11.5033

Tu, L., **Wang, J.**, **Abutalebi, J.**, **Jiang, B.**, **Pan, X.**, **Li, M.**, **Gao, W.**, **Yang, Y.**, **Liang, B.**, **Lu, Z.**, & **Huang, R**. (2015). Language exposure induced neuroplasticity in the bilingual brain: A follow-up fMRI study. *Cortex*, **64**, 8–19. https://doi.org/10.1016/j.cortex.2014.09.019

Wang, J., **Wang, X.**, **Xia, M.**, **Liao, X.**, **Evans, A.**, & **He, Y**. (2015). GRETNA: A graph theoretical network analysis toolbox for imaging connectomics. *Frontiers in Human Neuroscience*, **9**. https://doi.org/10.3389/fnhum.2015.00386

Wang, R., **Fan, X.**, **Liu, C.**, & **Cai, Z. G**. (2016). Cognitive control and word recognition speed influence the Stroop effect in bilinguals: Bilinguals stroop effect. *International Journal of Psychology*, **51**(2), 93–101. https://doi.org/10.1002/ijop.12115

Wang, S., **Allen, R. J.**, **Fang, S.-Y.**, & **Li, P**. (2017). Cross-modal working memory binding and L1–L2 word learning. *Memory & Cognition*, **45**(8), 1371–1383. https://doi.org/10.3758/s13421-017-0731-2

Wei, X., **Gunter, T. C.**, **Adamson, H.**, **Schwendemann, M.**, **Friederici, A. D.**, **Goucha, T.**, & **Anwander, A**. (2024). White matter plasticity during second language learning within and across hemispheres. *Proceedings of the National Academy of Sciences*, **121**(2), e2306286121. https://doi.org/10.1073/pnas.2306286121

Wu, J., **Yang, J.**, **Chen, M.**, **Li, S.**, **Zhang, Z.**, **Kang, C.**, **Ding, G.**, & **Guo, T**. (2019). Brain network reconfiguration for language and domain-general cognitive control in bilinguals. *NeuroImage*, **199**, 454–465. https://doi.org/10.1016/j.neuroimage.2019.06.022

Xia, M., **Wang, J.**, & **He, Y**. (2013). Brain net viewer: A network visualization tool for human brain connectomics. *PLoS ONE*, **8**(7), e68910. https://doi.org/10.1371/journal.pone.0068910

Yeh, F.-C., **Vettel, J. M.**, **Singh, A.**, **Poczos, B.**, **Grafton, S. T.**, **Erickson, K. I.**, **Tseng, W.-Y. I.**, & **Verstynen, T. D**. (2016). Quantifying differences and similarities in whole-brain white matter architecture using local connectome fingerprints. *PLOS Computational Biology*, **12**(11), e1005203. https://doi.org/10.1371/journal.pcbi.1005203

Zhang, J., **Li, H.**, **Zhang, M.**, **Wang, Z.**, **Ao, X.**, **Jian, J.**, **Wei, N.**, **Liu, H.**, **Ding, G.**, & **Meng, X**. (2023). Functional preference of the left inferior parietal lobule to second language reading. *NeuroImage*, **270**, 119989. https://doi.org/10.1016/j.neuroimage.2023.119989

Zhang, J., **Zhuang, L.**, **Jiang, J.**, **Yang, M.**, **Li, S.**, **Tang, X.**, **Ma, Y.**, **Liu, L.**, & **Ding, G**. (2022). Brain fingerprints along the language hierarchy. *Frontiers in Human Neuroscience*, **16**, 982905. https://doi.org/10.3389/fnhum.2022.982905

Zou, L., **Abutalebi, J.**, **Zinszer, B.**, **Yan, X.**, **Shu, H.**, **Peng, D.**, & **Ding, G**. (2012). Second language experience modulates functional brain network for the native language production in bimodal bilinguals. *NeuroImage*, **62**(3), 1367–1375. https://doi.org/10.1016/j.neuroimage.2012.05.062

Zuo, X.-N., & **Xing, X.-X**. (2014). Test-retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: A systems neuroscience perspective. *Neuroscience & Biobehavioral Reviews*, **45**, 100–118. https://doi.org/10.1016/j.neubiorev.2014.05.009

https://doi.org/10.1017/S1366728925000185 Published online by Cambridge University Press