Second language learning tunes the language control network: a longitudinal fMRI study

Cong Liu, Angela de Bruin, Lu Jiao, Zilong Li & Ruiming Wang

To cite this article: Cong Liu, Angela de Bruin, Lu Jiao, Zilong Li & Ruiming Wang (2020): Second language learning tunes the language control network: a longitudinal fMRI study, Language, Cognition and Neuroscience, DOI: 10.1080/23273798.2020.1856898

To link to this article: https://doi.org/10.1080/23273798.2020.1856898

Published online: 10 Dec 2020.
Second language learning tunes the language control network: a longitudinal fMRI study

Cong Liu a,b, Angela de Bruin c, Lu Jiao a, Zilong Li b and Ruiming Wang b

a Department of Psychology, Normal College & School of Teacher Education, Qingdao University, Qingdao, People’s Republic of China; b Key Laboratory of Brain, Cognition and Education Sciences, Ministry of Education, & Guangdong Provincial Key Laboratory of Mental Health and Cognitive Science, Center for Studies of Psychological Application, School of Psychology, South China Normal University, Guangzhou, People’s Republic of China; c Department of Psychology, University of York, York, United Kingdom

ABSTRACT
The current longitudinal study investigated how classroom second language (L2) learning modulates the neural correlates of bilingual language control during language production. Chinese college freshmen majoring in English undertook two test sessions (i.e. pre-learning and post-learning) over the course of one year. Specifically, while in the scanner, participants were instructed to name pictures in either their first or second language in response to cues. Behavioral results showed that language switch costs in the post-learning session were reduced as compared to the pre-learning session. fMRI results showed that, compared to the pre-learning session, the connectivity strength between the dorsal anterior cingulate gyrus and the left caudate was increased in the post-learning session. Critically, this increased connectivity strength was correlated with the reduction in language switch costs. These findings suggest that the language control network used during bilingual language production could be modulated by long-term L2 learning in a naturalistic classroom setting.

ARTICLE HISTORY
Received 14 October 2019 Accepted 20 November 2020

KEYWORDS
Second language learning; language production; language control; longitudinal design; fMRI

1. Introduction

Bilinguals activate both of their languages in parallel during speech production (e.g. Costa et al., 2000; Declerck, 2020). Therefore, to speak in the intended language, language control processes are implemented to minimise cross-language interference (for a review, see Declerck & Philipp, 2015).

One of the most common experimental tasks used to investigate the mechanisms underlying language control is the language switching task (e.g. Liu et al., 2019; Meuter & Allport, 1999; Timmer et al., 2017), in which participants are instructed to name digits or pictures in either their first language (L1) or second language (L2) in response to cues. This task typically shows that switching from one language to another (i.e. a switch trial) yields longer reaction times than repeating the same language in two subsequent trials (i.e. a non-switch trial). This switch cost is often used as a measure of language control (e.g. Meuter & Allport, 1999). A prominent model of language control, the Inhibitory Control Model (ICM, Green, 1998), proposes that inhibitory control is involved in language control during bilingual language production. Switch costs might originate from inhibition over the non-target language and the need to overcome this inhibition when switching to a new target language. In the present study, we set out to investigate how language control mechanisms are shaped by long-term classroom L2 learning over the course of one year.

Previous fMRI studies have identified a number of key brain regions involved in bilingual language control processes (Abutalebi & Green, 2007; Calabria et al., 2018). Abutalebi and Green’s (2007) Neurocognitive Language Control (NLC) model in particular highlighted that the language control network mainly includes neural regions such as (1) the dorsal anterior cingulate gyrus (dACC), (2) the left caudate nucleus (LCN), (3) the left inferior frontal gyrus (IFG), and (4) the left inferior parietal lobe (IPL). The dACC has been shown to be important in conflict monitoring and has been defined as the monitoring system for language control in bilinguals (Green & Abutalebi, 2013; Tu et al., 2015). Increased activation in the LCN has been found during language switching (De Bruin et al., 2014; Zou et al., 2012) and has been associated with language selection in the face of cross-language competition (Abutalebi & Green, 2008). The left IFG has been associated with domain-general inhibitory control and activation in the left IFG might decrease after a short amount of language switching training.
Notably, the Adaptive Control Hypothesis (ACH), proposed by Green and Abutalebi (2013), postulated that bilingual language control mechanisms used during language production are not fixed but adapt depending on the specific type of bilingual experience. Evidence supporting this argument mainly comes from studies using a cross-sectional design comparing different groups of bilinguals (e.g. Kousaie et al., 2017) or comparing bilinguals to monolinguals (e.g. Li et al., 2015). Given that cross-sectional designs might be hindered by individual differences within groups and might not directly reveal the causal relationship between bilingualism/language learning and language control mechanisms, recent studies have started to conduct longitudinal training studies to test the adaptability of language control mechanisms (Hervais-Adelman et al., 2015; Kang et al., 2017; Tu et al., 2015; Wu et al., 2017). For example, in one behavioural study, Wu et al. (2017) trained a group of unbalanced Chinese–English bilinguals on a cued picture-naming task four times on two consecutive days (two times per day) in the lab, and they found that switch costs in the L1 decreased with training. This suggested that a short language-switching training programme improved the efficiency of language control. Moreover, in one fMRI study, a group of unbalanced Chinese–English bilinguals completed an 8-day cued picture naming training in the lab. The results showed that the behavioural switch costs and activation of language control areas including the dACC and the LCN were reduced in the after-training session as compared to the before-training session (Kang et al., 2017). However, these training studies only tracked short-term training effects in a laboratory setting, which lacks ecological validity and might not reflect changes in language control mechanisms in daily life. In other words, it remains to be examined whether long-term extensive language training in a more naturalistic setting modulates language control mechanisms in bilinguals. The present study therefore examined if and how one year of L2 learning in a naturalistic classroom setting shapes language control mechanisms.

Following the Adaptive Control Hypothesis, the more recent Dynamic Restructuring Model (DRM) (Platsikas, 2020) proposed a time-course (i.e. initial exposure, consolidation and peak efficiency) for structural brain changes and suggested that these adaptations are dynamic and depend on the quantity and quality of the language learning and switching experience. For instance, Platsikas (2020) notes that during the early exposure to a new language grey matter changes seem to occur in brain regions related to language control (e.g. IFG, ACC and Caudate), and these changes are typically documented in non-immersed sequential bilinguals as well as in participants enrolled in intensive language training studies. Although the DRM mainly focused on structural neuroplasticity, it suggests that L2 learning experiences might modulate brain regions related to language control.

To date, research on the neural mechanisms underlying language control and the role of L2 language learning is still in its infancy (Abutalebi & Green, 2016; Liu et al., 2020). Only a few studies have attempted to investigate changes in brain plasticity over time in relation to classroom L2 learning, but they did not specifically explore how language control mechanisms adapt to classroom L2 learning (Grant et al., 2015; Legault et al., 2019). For instance, Grant et al. (2015) found that the connectivity between the caudate and the ACC in classroom L2 learners increased over the course of one academic year (i.e. 4–5 months). Moreover, another similar study found that English-Spanish language learners showed greater cortical thickness in the left ACC and right middle temporal gyrus (MTG) after four months of L2 learning as compared to controls (Legault et al., 2019). However, most of these studies have either looked at structural differences without using any behavioural measures of language control (DeLuca et al., 2019a; for reviews, see Li et al., 2014; Platsikas, 2020) or focused on language comprehension (such as the Grant et al. (2015) study, which looked at lexical competition in a homonym task). It thus remains an open question how long-term classroom L2 learning shapes language control networks, in particular during bilingual language production. The current study aims to address this gap by examining a group of classrooms English L2 learners across a period of approximately one year.

In the present study, a group of Chinese college freshmen (i.e. unbalanced Chinese-English bilinguals) who were about to start a major in English were recruited and performed the same cued picture naming task in the MRI scanner before and after one year of classroom L2 learning. During this year, the participants took various English courses, including oral, writing, and intensive reading courses in an all-English teaching and learning environment. This L2 learning experience reflects how L2 learning often happens in a naturalistic and immersive classroom in real-life. Testing L2 classroom learners in this setting provides good ecological validity to investigate how these experiences influence
language control mechanisms, focusing on the ability to switch between languages. According to the ACH, DRM and the findings from previous studies (Green & Abutalebi, 2013; Pliatsikas, 2020), we hypothesised that classroom L2 learning over the course of one year could modulate the neural mechanisms of language control in bilingual language production. We used language switching costs as a measurement of bilingual language control. Specifically, we predicted that the activation of brain areas related to language control would differ before and after the training, and that the functional connectivity between them would strengthen. Moreover, we assessed whether developmental changes in language switch costs correlated with neural changes in the bilingual language control network, which could indicate a direct link between language control and their corresponding neural bases while adapting to long-term classroom L2 learning experiences.

2. Method

2.1. Participants

Twenty-five Chinese-English bilinguals from South China Normal University took part in this study. Three participants were excluded due to not participating in the post-learning session, leaving 22 participants (20 females, mean age: 18.35 years, SD: 0.71) for the analysis. All of the participants were college freshmen majoring in English, and had started learning English on average at age 7.40 (SD = 2.81). They were all right-handed with normal or corrected-to-normal vision. All participants signed written informed consent before the experiment and got paid for their participation after the experiment. This research was approved by the Research Ethics Committee at South China Normal University.

In both the pre-learning and post-learning session, participants completed a language history questionnaire and rated their proficiency levels in both languages on a 7-point scale (1 = lowest level of proficiency, 7 = highest level of proficiency). Paired sample t-tests showed that these participants were unbalanced bilinguals with a higher proficiency level in Chinese than English (t (21) = 9.740, p < 0.001 in the pre-learning session; t (21) = 8.026, p < 0.001 in the post-learning session). Paired sample T-tests also showed that these participants’ self-rated proficiency scores in English improved significantly after one year of L2 learning in a classroom setting (see Table 1, t (21) = 2.856, p < 0.01). To confirm the reliability of these self-ratings, participants were asked to complete an Oxford Placement Test (OPT) in the post-learning session. The OPT contains 25 multiple choice questions and a cloze test, with a maximum score of 50, and higher scores indicated higher levels of English proficiency. This test has been widely used to measure language proficiency in previous language-production studies (e.g. Jiao et al., 2020; Yang et al., 2018). Pearson correlation analyses showed a correlation between the OPT score and the self-rated English proficiency score in the post-learning session (r = 0.355, p = 0.041), which (to some extent) confirmed the reliability of the self-rated proficiency data.

2.2. Procedure

In the present study, participants were tested in two sessions, once in October (i.e. pre-learning session) and once one year later (i.e. post-learning session). In both sessions, participants performed the same language switching task inside the MRI scanner.

For the language switching task, 16 black and white drawings were selected from the database of Zhang and Yang (2003), of which 4 pictures were used in the practice phase. According to the norming data from Zhang and Yang (2003) and Snodgrass and Vanderwart (1980), attributes such as familiarity, visual complexity, and image agreement are matched. Before the formal experiment, participants were familiarised with the pictures and their corresponding Chinese and English names until they correctly named all the pictures. Then they completed one practice session with 12 trials outside the scanner.

The language switching task included two runs. Each run contained 48 switch trials and 48 non-switch trials, which were presented in a pseudo-randomized manner within a rapid event-related design. Each trial began with a fixation cross together with a Chinese or American flag for 500 ms, followed by a picture for 1500 ms. Participants were instructed to name the picture in the language indicated by the Chinese or American flag as quickly and accurately as possible in a soft voice. The Chinese flag indicated that the picture had to be named in Chinese while the American flag indicated it had to be named in English. Then a fixation cross was presented during a jittered inter-stimulus-interval (ISI) lasting between 2000ms and

Table 1. Means (and SDs) of the language proficiency self-ratings for both Chinese and English.

<table>
<thead>
<tr>
<th></th>
<th>Pre-learning</th>
<th>Post-learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese*</td>
<td>6.05 (0.87)</td>
<td>5.93 (0.60)</td>
</tr>
<tr>
<td>English**</td>
<td>4.25 (0.77)</td>
<td>4.76 (0.53)</td>
</tr>
<tr>
<td>OPT score</td>
<td>41.5 (2.93)</td>
<td></td>
</tr>
</tbody>
</table>

*Self-rated proficiency score from 1 (lowest) to 7 (highest). ** p < 0.01.
5500 ms (in steps of 500 ms). The timing and order of trial presentation within each run was optimised for estimation efficiency using optseq2 (http://surfer.nmr.mgh.harvard.edu/optseq/). Each run lasted approximately 8 min (see Figure 1 for an example of trials sequence). Because of technical limitations, we re-collected the verbal response data in a behavioural test after the scanning session.

In addition, participants completed a color-shape switching task to measure cognitive flexibility and a modified flanker task to measure inhibitory ability, in both the pre-learning and post-learning sessions. We included them in the current manuscript to assess whether behavioural changes between sessions are uniquely associated with the language switching task. The color-shape switching task included one run with 128 trials (containing 64 switch trials and 64 non-switch trials). Each trial began with a fixation cross together with a rainbow or geometric figure for 500 ms, followed by a stimulus for 1500 ms. Participants were instructed to decide if the stimulus was red or yellow when the cue was a rainbow and to decide if the stimulus was a circle or triangle when the cue was a geometric figure. Participants indicated their responses via a left or right button press. Afterwards, a fixation cross was presented during a jittered ISI lasting between 2000ms and 5500 ms (in steps of 500 ms). The modified flanker task included two runs and each run contained 24 neutral trials, 24 congruent trials, 24 incongruent trials and 24 no-go trials. On each trial, a fixation cross was first presented during a jittered ISI lasting between 2000ms and 5500 ms (in steps of 500 ms). Then an array of five stimuli including a central arrow and two stimuli on either side of it (i.e. flankers) was present for 2000ms. Participants were instructed to respond by pressing the left button when the central arrow pointed to the left and by pressing the right button when it pointed to the right. On neutral trials, the flankers were diamonds, which were not associated with any response. On congruent trials, the flankers were arrows pointing in the same direction as the target. On incongruent trials, the flankers were arrows pointing in the opposite direction of the target. On no-go trials, the flankers were Xs, which indicated that subjects should withhold their response.

2.3. MRI data acquisition

MRI data were collected using a 3 T Siemens Trio scanner with 12-channel phase array head coil at the MRI centre of South China Normal University. Functional images were acquired using T2-weighted gradient-echo planner imaging (EPI) sequence with the following parameters: TR = 2000ms, TE = 30 ms, flip angle = 90°, FOV = 204x204 mm², matrix = 64x64, slice thickness = 3.5 mm, voxel size = 3 x 3 x 3.5 mm³. In addition, high-resolution T1-weighted anatomical images were acquired using the MPRAGE sequence with the following parameters: TR = 1900ms, TE = 2.52 ms, flip angle = 9°, FOV = 256 x 256 mm², matrix = 204 x 204, slice thickness = 1 mm, voxel size = 1 x 1 x 1 mm³.

2.4. Data analysis

2.4.1. Preprocessing

MRI images were preprocessed using Statistical Parametric Mapping (SPM 12; www.fil.ion.ucl.ac.uk) implemented in Matlab R2015a (Mathworks Inc., Sherborn, MA). Imaging data were realigned, slice time corrected, normalised to Montreal Neurological Institute (MNI) space, resliced to a voxel size of 3 mm³, and smoothed with 8 mm Gaussian kernel. Artifact/outlier scans were excluded using the Artifact Detection Tools (ART, www.nitrc.org/projects/artifact_detect). An image was defined as an outlier if the average intensity deviated 3 SDs from the mean intensity in the session or if the composite head movement exceeded 1 mm from the previous image.

2.4.2. Whole brain analyses

Statistical analyses were performed by modelling different conditions on a voxel-by-voxel basis using the
General Linear Model. In the first level analysis, two types of effects of interest were assessed by computing the following statistical contrasts: (a) Switch trials versus Non-switch trials in the pre-learning session; (b) Switch trials versus Non-switch trials in the post-learning session. Movement parameter estimates produced by the realignment procedure were entered as covariates of no interest in order to correct for potential movement artifacts. In the second level analysis, one-sample t tests were first performed for the two aforementioned contrasts. Then, to examine potential L2 learning effects on switch costs, the neural activation of the switch cost (i.e. the difference between Switch trials and Non-switch trials) was compared between the pre-learning and post-learning sessions using a paired t-test. Only activations containing at least 100 contiguous voxels within a cluster (p < 0.05, FDR corrected) were reported as significant.

2.4.3. ROI analyses
As indicated by the Neurocognitive Language Control (NLC) model (Abutalebi & Green, 2007), the language control network includes the dACC, LCN, left IFG and left IPL. We chose four seeds based on the NLC model and the ROI coordinates were derived from three previous studies: the LCN and left IFG from Luk et al. (2012), a meta-analysis study on neural regions involved in bilingual language control; the dACC from Abutalebi et al. (2012); and the left IPL from Barbeau et al. (2017). These ROI coordinates were also chosen in a recent study by Gullifer et al. (2018), which suggested that both static components of language acquisition (e.g. age of acquisition) and the social diversity of language use contribute to adaptive changes in brain networks involved in bilingual language control, so we chose them as reference for our ROIs. In addition, two neural regions associated with language processing, namely the bilateral posterior superior temporal gyrus (pSTG), were selected as control regions in which no changes were expected (Friederici, 2011), and their ROI coordinates were derived from Sulpizio et al. (2020). All ROIs were 6 mm spheres and their coordinates are represented in MNI space (see Table 2 for seed coordinates). Beta values were obtained from the single-subject contrast images (non-switch trials vs. switch trials) and were exported for group-level analyses. For each ROI’s beta value, we performed a paired sample t test comparing switch costs between the pre-learning and post-learning session.

2.4.4. Connectivity analyses
ROI-to-ROI functional analysis was performed with the CONN toolbox (v. 18b: https://www.nitrc.org/projects/conn/; Whitfield-Gabrieli & Nieto-Castanon, 2012) with SPM12. The CONN toolbox has been used for functional connectivity analyses in event-related designs (see Beaty et al., 2015; Berger et al., 2018), by conducting a generalised psycho-physiological interaction (gPPI) approach (see CONN manual at https://web.conn-toolbox.org/resources/manuals/#h.p_aiqzwnbuSuss). To correct for confounds of physiological noise and motion (Chai et al., 2012), the CONN toolbox implemented the anatomical component-based noise correction method (CompCor; Behzadi et al., 2007) identifying principal components associated with the segmented white matter and cerebrospinal fluid. These components were entered as confound regressors along with realignment parameters in the first-level analysis. In addition, in order to exclude simple task-related activation effects, the main effects of the task conditions (switch, non-switch) were included as confound regressors in our connectivity analysis. A temporal bandpass filter (0.01–inf. Hz) was applied to the time series. Within the gPPI approach, the mean time-series, averaged across all voxels within each seed was correlated with each other by bivariate correlations. The final results from group-level statistics provided Fisher-z transformed correlation coefficient values. Paired sample t-tests were used to examine whether network connectivity changed over the one-year period. FDR corrections for multiple comparisons were applied at an adjusted significance level of p < .05.

2.4.5. Correlation analyses
Correlation analyses were conducted to assess the correlation between changes in behavioural language switch costs (i.e. RTs) and significant changes in brain connectivity (i.e. Fisher’s z values reflecting functional connectivity across ROIs). We also carried out correlational analyses between significant brain connectivity changes and behavioural changes on the flanker and color-shape switching tasks.

3. Results

3.1. Behavioural results
For the language switching task, RTs over 2.5 SDs from the mean in each condition were excluded from
the behavioural analysis (cf. Liu et al., 2019). We only performed statistical analyses on RTs, as accuracy was generally high for all participants (> 95%) (see Figure 2).

Analyses were conducted using linear mixed-effects models with crossed random effects for participants and items using the lme4 package (Bates et al., 2014) and the lmerTest package (Kuznetsova et al., 2014) in R (version 3.4.4). We used a mixed-effects model because it allows random effects of participants and items to be considered simultaneously, making the results generalisable to other subjects and items.

We fit a mixed-effect model for log RT data, with test session (pre-learning vs. post-learning), trial type (non-switch vs. switch), and their interaction as fixed effects. All variables were coded using mean-centered contrast coding (i.e. pre-learning = −0.5, post-learning = 0.5; non-switch = −0.5, switch = 0.5), and the results thus present main effects analogous to ANOVAs. The set of pictures was repeated frequently. To control for potential picture repetition effects, the model also included item repetition as a continuous variable, as well as the interactions with the other two variables. To reduce collinearity, the continuous fixed effect (i.e. item repetition) was z-scored. After removal of correlations between the random slopes and the random intercepts (“no random correlations”, Barr et al., 2013), the final converging model included intercepts with a maximal random effects structure (i.e. participant and item slopes for test session, trial type, item repetition and their interactions). The final model was checked for collinearity between variables through VIFs (Frank, 2011), and VIFs were below 2.5 (De Bruin et al., 2018).

As shown in Table 3, the effect of test session was not significant (t = −1.44, p = 0.165), suggesting that overall RTs were similar in the pre- and post-learning session. There was a significant effect of trial type (t = 3.31, p = 0.007), indicating that switch trials (757 ms) were named more slowly than non-switch trials (734 ms). Critically, the interaction between test session and trial type reached significance (t = −2.14, p = 0.033), suggesting that the switch cost in the post-learning session (11 ms) was smaller than in the pre-learning session (36 ms). However, there was no significant effect of item repetition (t = 1.53, p = 0.142) and no significant interactions between Trial type × Item repetition or between Test session × Trial type × Item repetition (ps > 0.05), indicating that item repetition did not affect overall RTs or the switch costs (which is of main interest here).

We also assessed whether there were behavioural changes in the color-shape switching and flanker tasks between sessions. Considering that it was not necessary to control the random effects of items in the color-shape switching task and flanker task, we conducted paired sample T-tests for both tasks. The results showed that the task switching costs (i.e. cognitive flexibility = switch trials – non-switch trials) in the color-shape switching task reduced significantly in the post-learning session as compared to the pre-learning session (82 ms vs. 50 ms; t (21) = 2.293, p = 0.032). However, the flanker effect (i.e. inhibitory ability = incongruent trials – congruent trials) remained the same across the pre- and post-learning sessions (110 ms vs. 111 ms; t (21) = −0.079, p = 0.938).

To explore the relationship between the language switching tasks and non-language tasks, we correlated the language switching costs with the task switching costs and the flanker effect in the pre- and post-learning sessions separately. The results showed no significant correlations between the language switching costs and task switching costs (r = −0.221, p = 0.324 in pre-learning session; r = 0.012, p = 0.958 in post-learning session), or between the language switching costs and flanker effect (r = 0.257, p = 0.248 in pre-learning session; r = 0.190, p =

![Figure 2](image_url). The mean RTs for (A) non-switch trials, switch trials and (B) language switch costs (switch trials – non-switch trials) in the pre-learning and post-learning session. Error bars show one standard error. **p < 0.01.
Table 3. Results for mixed-effects model on log RTs.

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Estimate</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>6.58</td>
<td>0.02</td>
<td>272.39</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Test session</td>
<td>-0.04</td>
<td>0.03</td>
<td>-1.44</td>
<td>0.165</td>
</tr>
<tr>
<td>Trial type</td>
<td>0.03</td>
<td>0.01</td>
<td>3.31</td>
<td>0.007</td>
</tr>
<tr>
<td>Item repetition</td>
<td>0.01</td>
<td>0.01</td>
<td>1.53</td>
<td>0.142</td>
</tr>
<tr>
<td>Test session × Trial type</td>
<td>-0.02</td>
<td>0.01</td>
<td>-2.14</td>
<td>0.033</td>
</tr>
<tr>
<td>Test session × Item repetition</td>
<td>0.00</td>
<td>0.01</td>
<td>0.21</td>
<td>0.832</td>
</tr>
<tr>
<td>Trial type × Item repetition</td>
<td>0.00</td>
<td>0.01</td>
<td>-0.06</td>
<td>0.951</td>
</tr>
</tbody>
</table>

0.397 in post-learning session). In addition, neither the correlation between the amount of change in the language switch costs and the amount of change in the non-language switching costs ($r = 0.049, p = 0.828$), nor the correlation between the amount of change in the language switch costs and the amount of change in the flanker effect reached significant ($r = 0.173, p = 0.440$).

3.2. fMRI results

3.2.1. Whole brain results

One-sample t-tests showed that on language switch compared to non-switch trials, the left middle frontal gyrus (MNI = $[-39, 45, 0]$, $t = 6.10$) and left IPL (MNI = $[-48, -39, 42]$, $t = 4.83$) were activated in the pre-learning session. There were no significant differences between switch and non-switch trials in the post-learning session. Critically, the paired sample t-test did not indicate significant differences in any brain region when we contrasted the neural activity for switch costs in the post-learning with that in the pre-learning (see Figure 3).

3.2.2. ROI results

The results of paired sample t-tests showed that, within the language control network, there was a significant decrease in switch costs between pre- and post-learning sessions in the beta values of the left IPL ($t = 2.668, p = 0.014$) but not the dACC ($t = 1.903, p = 0.071$), LCN ($t = 0.572, p = 0.573$) and left IFG ($t = 1.045, p = 0.308$). However, none of the regions survived FDR corrections ($p > 0.0125$) (see top panel of Figure 4). In addition, we included ROIs related to language processing that are not part of the language control network. There was no significant decrease in the beta values of these control ROIs between pre- and post-learning sessions: left pSTG ($t = 0.662, p = 0.515$) and right pSTG ($t = 1.045, p = 0.308$) (see bottom panel of Figure 4).

3.2.3. Connectivity results

As shown in Figure 5A, the results of the ROI-to-ROI connectivity analysis indicated that, for the language control network, only the connectivity strength between the dACC and LCN was significantly increased in the post-learning as compared to the pre-learning session ($t = 3.4, p = 0.008$, FDR corrected). There were no significant changes in connections between other areas of the language control network, nor between regions associated with language processing ($p > 0.05$).

3.2.4. Correlation results

In an additional correlational analysis, we found that the increase in connectivity between the dACC and LCN was significantly correlated with the reduction in the behavioural language switching costs ($r = -0.502, p = 0.017$). Specifically, the stronger the increase in connectivity between the dACC and LCN, the smaller the language switching cost was in the post-learning compared to the pre-learning session (see Figure 5B). No significant correlations were observed between the increase in dACC-LCN connectivity and behavioural changes in the non-language switching task ($r = -0.060, p = 0.789$) or the flanker task ($r = -0.201, p = 0.370$)

4. Discussion

In the present study, we investigated how long-term classroom L2 learning shapes the neural mechanisms of bilingual language control in language production. Chinese College freshmen majoring in English completed a language switching task in the MRI scanner before and after classroom L2 learning over the course of one year. Behavioural results showed that the language switch cost in the post-learning session was significantly smaller than the cost in the pre-learning session. Furthermore, fMRI data revealed that the connectivity between the dACC and LCN significantly increased with classroom L2 learning, and this increase was significantly correlated with the improvements in the behavioural language switching cost. These results
suggest that classroom L2 learning induces adaptive changes in brain networks involved in bilingual language control.

4.1. Developmental changes in behaviour

The Adaptive Control Hypothesis (ACH) and Dynamic Restructuring Model (DRM) proposed that language control mechanisms during bilingual language production adapt to specific types of language experience (Green & Abutalebi, 2013; Pliatsikas, 2020). The present study explored the performance of Chinese College freshmen majoring in English (i.e. Chinese-English bilinguals) and observed that they showed decreased language switch costs after classroom L2 learning over the course of one year.

This finding is consistent with previous studies (Kang et al., 2017; Wu et al., 2017) that showed reduced language switch costs induced by short-term language-switching training, which was interpreted as indicating improved language control efficiency. While previous results have demonstrated short-term training effects on language control in laboratory settings, the current study reveals a long-term language training effect in a more naturalistic classroom setting. The current study showed that language switch costs

Figure 4. The beta values reflecting the language switch cost in the pre-learning and post-learning session in ROIs from the language control network (top panel) and ROIs from the language processing network (bottom panel).

Figure 5. Connectivity between the dACC and LCN correlated with language switching costs. (A) Left panel indicating that the connectivity strength between the dACC and LCN was significantly increased in the post-learning as compared to the pre-learning session. (B) Association between connectivity (Fisher’s z) and behavioural performance on the language switching task.
decreased after classroom L2 learning over the course of one year, which may reflect improved language control efficiency.

Behavioural changes after L2 learning were not just observed on the language switching task but also on the color-shape switching task. This suggests that classroom L2 learning can affect switching in general, both in language and non-language switching tasks. However, the present study showed no significant correlation between the language switching costs and task switching costs. This could suggest that classroom L2 learning could affect language and task switching without a direct overlap between the two types of switching. The flanker effect in the modified flanker task remained the same across the two sessions, suggesting that non-verbal inhibition was not influenced by classroom L2 learning. This null effect could furthermore suggest that the observed changes in the language switching task and color-shape switching task were not purely the result of task practice or training effects.

4.2. Changes in functional connectivity in the bilingual language control network

The main finding from the neuroimaging data in the current study was that the connectivity strength between the dACC and LCN was significantly higher in the post-learning as compared to the pre-learning session. This finding is consistent with previous work indicating that the dACC and LCN are two key brain regions in the language control network (Garbin et al., 2011; Kang et al., 2017; Zou et al., 2012). The dACC has been reported to be responsible for domain-general conflict monitoring (Botvinick et al., 2001; Botvinick et al., 1999), and the LCN might be responsible for inhibiting cross-language interference during language selection in bilinguals (Abutalebi et al., 2013; Abutalebi & Green, 2008; Branzi et al., 2015). Both regions have been associated with language switching in bilinguals (Abutalebi & Green, 2007; Branzi et al., 2015; Guo et al., 2011; Hosoda et al., 2012; Zou et al., 2012). The links between these two regions may reflect that they work together to monitor the language environment and to resolve competition between the languages. Thus, in the present study, the increased connectivity strength between the dACC and LCN might suggest that the participants’ ability to monitor and resolve conflict (i.e. language control abilities) improved after long-term L2 learning in a classroom setting. It should be emphasised that this concerns an increase in functional connectivity. In contrast, previous studies focusing on structural connectivity have observed that language learning decreased isotropic diffusivity in white matter, which signified more efficient communication between frontal, parietal, and subcortical language-related regions (Platsikas, 2020). There were no significant changes in neural activation in brain regions associated with language processing outside the language control network. Moreover, these regions revealed no significant changes in functional connectivity. The observed changes thus appear to be associated with the language control network specifically.

Critically, a significant negative correlation between the behavioural improvements in the language switching task (i.e. reduced language switch cost) and neural changes in functional connectivity was observed, suggesting that the increased connectivity between the dACC and LCN was related to enhanced language switching performance.

These findings show how language control mechanisms can be shaped by long-term classroom L2 learning. This is in line with Grant et al. (2015), who showed increased connectivity between the caudate and the ACC after classroom L2 learning over the course of one academic year. While some previous studies have shown that L2 learning might be associated with functional changes during language comprehension (Grant et al., 2015; Legault et al., 2019), our study focuses on language production. By combining neural and behavioural data, our study indicates that language control used during language production is shaped by classroom L2 learning.

In our ROI analysis, the neural activation of some brain regions in the language control network (e.g. left IPL) significantly decreased from the pre-learning session to the post-learning session, which was in line with some previous studies showing decreased activation in language control brain regions such as the dACC and LCN with short-term language switching training (Kang et al., 2017; but see Grant et al., 2015). This pattern of decreased activation may thus reflect an optimisation of the language control system within the L2 classroom learning environment. However, while there was a change in connectivity strength, no significant decrease in neural activation was observed in the dACC and LCN in the present study, although there was a trending decrease in both regions. This differs from the findings in Kang et al. (2017). We speculate that such difference might arise from the different training types in both studies. While the training in Kang et al. (2017) was an intensive short-term language switching training in a laboratory setting, the long-term L2 learning in the current study was in a more naturalistic classroom setting. These different training types may eventually lead to different neuroplastic changes in areas responsible for language control. In our whole
4.3. Limitations

Our fMRI study did not include a control group and it could be argued that the decrease in switch costs was related to task-related training. To address this limitation, we added a behavioural control group. Specifically, 25 college freshmen who were not majoring in foreign languages (22 females, mean age: 19.32 years, SD: 0.75; Self-rated Chinese proficiency = 5.94, Self-rated English proficiency = 4.09, t = 9.323, p < 0.001) were recruited to behaviourally complete the language switching task twice about one month apart. The results showed that the language switch cost remained the same between the pre- and post-learning sessions (12 ms vs. 9 ms, t = −0.54, p > 0.05). Together with the finding that changes in connectivity were only observed in language control regions, this suggests that the observed changes were specifically related to L2 learning affecting language control.

While our results indicate that classroom L2 learning can shape language control, they do not show which aspects of L2 learning are most important. The use of a naturalistic classroom setting is important for ecological validity (see Blanco-Elorrieta et al., 2018; DeLuca et al., 2019b). Future studies should use these environments to assess which aspects of L2 learning contribute to changes in language control. In addition, our study mainly included female participants (i.e. 20/22 participants were female), which might constrain the generalisation of the findings. Future studies should use more balanced designs with similar numbers of female and male participants.

5. Conclusion

The current study examined behavioural and neural changes in language control before and after one year of English classroom learning. Our results show that L2 learning may lead to increases in connectivity strength between the dACC and LCN, two key brain regions involved in language control that have been implicated in previous studies (Garbin et al., 2011; Kang et al., 2017). This increased connectivity strength was correlated with the reduction in behavioural RT language switch costs. These findings provide considerable support to the idea that long periods of L2 classroom learning may induce significant neuroplastic changes in areas responsible for language control. These findings corroborate recent theoretical perspectives on the adaptive features of bilingual language control mechanisms, including the ACH (Green & Abutalebi, 2013) and DRM (Pliatsikas, 2020). As postulated in both theories, bilinguals’ cognitive and language control processes are shaped by language learning and switching experiences. Combining the findings from previous studies and the current findings, we show the flexibility and plasticity of bilingual language control mechanisms, which not only adapt to short-term language-switching training, but also as the result of long-term extensive language training in naturalistic classroom settings.

Note

1. We did not include “language” in the linear mixed-effects models analyses because there were no interactions between language and other variables.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by National Social Science Foundation of China [Grant Number 19AYY009].

ORCID

Cong Liu http://orcid.org/0000-0001-8459-2637

References


