How bilingual experience and executive control influence development in language control among bilingual children

Citation for published version:

Digital Object Identifier (DOI):
10.1111/desc.12865

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Developmental Science

Publisher Rights Statement:
This is the peer reviewed version of the following article:
This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Title Page

(i) Title: How bilingual experience and executive control influence development in language control among bilingual children

(ii) Running title: influence of executive control on language control

(iii & iv) Authors:

MAKI KUBOTA
School of Philosophy, Psychology, and Language Sciences, University of Edinburgh

NICOLAS CHEVALIER
School of Philosophy, Psychology, and Language Sciences, University of Edinburgh

ANTONELLA SORACE
School of Philosophy, Psychology, and Language Sciences, University of Edinburgh

(v) We have no conflict of interest to declare.

(vi) We would first and foremost like to thank the parents and the children who participated in the research. Special thanks go to Japan Overseas Education Services (JOES) as well as Setagaya Children’s Center for assisting us with participant recruitment and test administration. We would also like to express our appreciation to the following foundations for providing financial support: Great Britain Sasakawa Foundation (grant number 5038), Japan Foundation Endowment Committee (grant number 5780416), and the British Association for Japanese Studies (award number JC2018-6-1).

(vi) The data that support the findings of this study are available on request from the corresponding author: Maki Kubota. The data are not publicly available due to privacy or ethical restrictions.
**Research Highlights**

- Bilingual children’s development in the L2 single trials which involves language-specific processing was predicted by L2 exposure and rate of improvement in executive control.

- Development in the L2 repetition trials (in the mixed block) which involves domain-general processing was predicted by the rate of improvement in executive control.

- Taken together, the findings suggest that language control development in bilingual children is modulated by both bilingual experience and improvement in executive control.
Abstract

This longitudinal study investigates whether the development in executive control and bilingual experience predicts change in language control in bilingual children. Children were tested twice over the course of one year, using the language-switching paradigm and the Simon task. The participants were Japanese-English bilingual ‘returnee’ children (ages 7–13), who returned to their first language (L1) environment after spending some years in a second language (L2) dominant environment. Testing these children upon their return to the L1 environment allowed us to disentangle the effect of age from bilingual experience, as they experienced an increase in age but a decrease in L2 exposure over time. Children who had less L2 exposure showed smaller improvement in baseline performance when naming pictures in English (i.e., when English was relevant across all trials). Moreover, development in trials where children had to switch between languages were modulated by development in executive control. That is, children who increased their performance in the English mixed repetition trials also performed better on the executive control task over time. Thus, development in executive control modulated change in language control among bilingual children, suggesting a positive relationship between language control and executive control in children’s development.

Keywords:
Bilingual returnee children, language control, bilingual experience, executive control, language-switching, longitudinal study
1. Introduction

Bilinguals activate both of their languages even when confining their speech to the target language (Grosjean, 1989). One would then expect that there is constant interference from one language to the other. However, bilinguals efficiently keep the two languages separate, successfully attending to one language while resisting interference from the other—a phenomenon often referred to as language control. Theoretical accounts of bilingual language control share the idea that this phenomenon is mediated by domain-general executive control (Craik & Bialystok, 2006; Green, 1998; Green & Abutalebi, 2013). However, the link between the two control mechanisms have not been consistently found in empirical studies, and debates still persist as to what extent they overlap. Furthermore, no study has yet examined how language control and executive control dynamically interact as they develop during childhood. The current study investigates the extent of overlap between the two control mechanisms and how bilingual experience influences language control over time in bilingual children’s development.

The relationship between language control and executive control has been widely investigated using the switching paradigm (Meuter & Allport, 1999), in which the bilinguals are required to switch between languages as well as non-linguistic tasks. Both paradigms are set to measure mainly two costs: switch and mixing costs. Switch costs refer to the difference between switch and repetition trials in which the bilinguals have to switch to the other language/task or repeat the same language/task within the same block. Mixing costs are the difference between repetition trials in the mixed language/task block and repetition trials in the single language/task block (i.e., single stimuli are tested individually). Although debates persist in the literature as to what the sources of these costs are, switch costs are often referred to as transient control processes whereas mixing costs reflect more sustained aspects.
of control (Bobb & Wodniecka, 2013). A number of studies have found that the magnitude of switch and mixing costs in language-switching correlate with their counterparts in task-switching (de Bruin, Roelofs, Dijkstra, & Fitzpatrick, 2014; Declerck, Grainer, Koch, & Philipp, 2017; Linck, Schwieter, & Sunderman, 2012; Liu, Fan, Rossi, Yao, & Chen, 2016; Liu, Rossi, Zhou, & Chen, 2014; Prior & Gollan, 2011;), while others found weak or no associations (Branzi, Calabira, Boscarino, & Costa, 2016; Calabria, Branzi, Marne, Hernández, & Costa, 2015; Jylkkä, Lehtonen, Lindholm, Kuusakoski, & Laine, 2018; Weissberger, Wierenga, Bondi, & Gollan, 2012).

It is widely acknowledged that language control is dependent on language-specific processes: language production involves a lexical selection process, in which the speakers must regulate the activation of lexical representations in order to choose the correct item that matches the intended meaning (Costa & Santesteban, 2004; Levelt, Roelofs, & Meyer, 1999). For example, when the intended meaning is the common four-legged domestic animal ‘dog’, speakers must choose the appropriate lemma dog out of other semantically related (and therefore activated) lemmas such as cat and tail—and this process occurs within the target language of a bilingual. Research based on the connectionist framework suggests that increased practice and use of the language creates stronger links between forms and concepts (Dijkstra, 2005; Kroll & Stewart, 1994; Michael & Gollan, 2005). Since bilinguals have less practice in each language than their respective monolinguals, their associative networks between words and concepts are weaker (i.e., “weaker links” hypothesis: Gollan & Acenas, 2004; Gollan & Silverberg, 2001; Gollan, Montoya, Fennema-Notestine, & Morris, 2005), hence potentially explaining why bilinguals are slower than monolinguals at naming pictures even in their dominant language (Gollan et al., 2005; Ivanova & Costa, 2008; Yan & Nicoladis, 2009). Thus, language-specific processes are dependent on bilingual experience
(specifically language exposure), and so bilingual experience should play a crucial role in language control.

Not only does bilingual experience influence language control, but it also enhances executive control, as observed in research comparing executive control in bilinguals to monolinguals (see Bialystok, 2009, for review). In addition to the task-switching paradigm, a commonly used task to measure executive control is the Simon task, which measures how well one can inhibit unwanted responses by manipulating the position of the stimulus and the position of the button in which they are instructed to press. In congruent trials, the position of the stimulus matches the position of the button, while the positions of the stimulus and the button do not match up in incongruent trials. The Simon effect is the difference in response time between congruent and incongruent trials. Unlike the Simon effect, which reflects inhibition per se, global response time (RT) (RTs of both congruent and incongruent trials) may reflect the ability to monitor competing sets of cues and selectively attend to the target cue (Martin-Rhee & Bialystok, 2008). Bilinguals outperform monolinguals on the Simon task as well as several other executive control tasks, and positive correlations have been reported between performance on the Simon task and language-switching (de Bruin et al., 2014; Linck et al. 2012). Although these findings speak to a ‘bilingual advantage’ in terms of executive control, direct comparison between language-switching and task-switching performance has generally found no or only weak correlations (Calabria, Hernandez, Branzi, & Costa, 2012; Klecha, 2013; Prior & Gollan, 2011), contributing to persisting debates on the reality of such a ‘bilingual advantage’ (for further discussion see Paap, Johnson, & Sawi, 2015).

This debate has also been approached developmentally by directly comparing the age-related changes in both language control and executive control (Calabria et al., 2015; Weissberger et al., 2012). Evidence for greater change across the life span in task-switching
than language-switching suggests that language control may not be as affected by age-related
decline as executive control (Calabria et al., 2015). In a similar vein, switch and mixing costs
vary differently as a function of age depending on whether they involved switching between
languages or non-linguistic tasks (Weissberger et al., 2012). Although such findings suggest
that language control and executive control follow different developmental trajectories,
bilingual experience was not carefully controlled in these studies, as they used self-rated
proficiency scores and/or did not control for other bilingual experiences such as language use
and exposure in each language. Therefore, it may be premature to draw any definite
conclusions about potential differences between the developments of language control and
executive control. For instance, smaller differences between older and younger bilinguals for
language control than executive control may be confounded by variations in bilingual
experience (exposure and use) with age.

The relationship between language control and executive control is also investigated
through potential language dominance effects on the magnitude of switch and mixing costs.
According to the Inhibitory Control Model (Green, 1998), switching from the less-dominant
to the dominant language in unbalanced bilinguals requires more time than from the
dominant to the less-dominant language, because bilinguals need to apply stronger inhibition
to the dominant language, and subsequently resolving that inhibition takes more effort and
time. Indeed, bilinguals have larger switch and mixing costs in their L1 or dominant-language
than L2 (Costa, Santesteban, & Ivanova, 2006; Linck et al., 2012; Meuter & Allport, 1999;
Philipp, Gade, & Koch, 2007; Schweiter & Sunderman, 2008). However, prior research has
mainly focused on younger to older adults, and only a number of studies examined this
phenomenon in bilingual children.
In addition, it is still unclear how language control and executive control interact developmentally. If the processes underlying these two control types overlap, then children who experience greater development in executive control should also do so in language control. What makes such investigation complex is that language control is largely influenced by bilingual experience (Bobb & Wodniecka, 2013). While executive control increases with age (for a review see Best & Miller, 2010), the bilingual experience is a fluid process, and does not necessarily increase in a linear manner with age. Some children may experience reduction in language exposure while others undergo increase in language input due to environmental or circumstantial changes. Thus, when examining the development of language control in children, the characteristics of the bilingual experience (in addition to executive control) must be taken into consideration.

Taken together, the development of language control, executive control, and bilingual experience appear deeply entangled with one another, especially among bilingual children as they often experience increase in both executive control and bilingual experience (Kohnert, Bates & Hernandez, 1999; Weissberger, Gollan, Bondi, Clark, & Wierenga, 2015). In contrast, executive control and bilingual experience can be more easily distinguished in children who lived in a L2 dominant environment for some years and returned to their L1 environment (“returnee” children), as these children, upon their return to the L1 environment, experience reduced input in their L2. Therefore, the present study strategically targets this specific population to further examine how development in language control specifically relates to bilingual experience and executive control. This population enables us to examine how language control changes as children experience reduction in (current) bilingual exposure, a major departure from the ‘traditional’ approach focusing on how increase in bilingual exposure affects language control.
Specifically, we tracked language control and executive control abilities of Japanese-English returnee children over the course of one year, from the point when these children had returned to their L1 Japanese environment. By that time, the children’s language exposure changed dramatically from having exposure to two languages (Japanese and English) to mainly receiving only Japanese input. We tested children at two time points using the language-switching task to examine how language-switching performance changes over time and, more specifically, whether the observed changes could be explained by children’s bilingual experience, executive control, or both. If language control is influenced by bilingual experience (in our current study, we used exposure as a measure of bilingual experience), then language exposure should predict the change in bilinguals’ language control. Specifically, we predicted that children who received less English exposure in Japan (thus less bilingual exposure) would experience a smaller degree of improvement in language control (as measured by language switch and mixing costs) over time. Alternatively, if executive control also plays a role in the development of language control, then the bilinguals with greater performance improvement in the non-linguistic interference task should also show greater progress on the language-switching task.

2. Methodology

2.1 Participants

The participants were 36 Japanese-English bilingual returnee children. An additional two children dropped out in the second round of testing and were thus removed from the final sample. The participant information is provided in Table 1. All participants had acquired Japanese as their L1 and English as L2. The parents of the bilingual children were all native speakers of Japanese. They all had similar socioeconomic status—all the mothers of the
children had a bachelor or a post-graduate degree. Prior to moving away from Japan, the children had very limited exposure to English. All the parents reported that their children began acquiring English through daily exposure to English in a foreign environment. Children attended schools with English as a medium of instruction while abroad. Half of the participants lived in a country where English is the official language (e.g., USA, UK, Canada), while the other half lived in a country where the official language is not English (e.g., China, France, Germany) but still received education in English. Although the latter group was exposed to a third language other than Japanese and English, none of the parents reported that their children could actually hold a conversation in the third language. The first round of testing took place when the participants had just returned to Japan. The average time elapsed between their return to Japan and the first round of testing (i.e., incubation period) was three months (SD = 1.0). Upon their return to Japan, all participants attended Japanese schools that operated under the curriculum set by the Japanese Ministry of Education.

Table 1: Summary of participant information

<table>
<thead>
<tr>
<th></th>
<th>Age of L2 acquisition</th>
<th>Length of residence in L2</th>
<th>Age at first round</th>
<th>Age at second round</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>5.0</td>
<td>4.0</td>
<td>9.8</td>
<td>10.8</td>
</tr>
<tr>
<td>SD</td>
<td>2.5</td>
<td>2.0</td>
<td>1.42</td>
<td>1.42</td>
</tr>
</tbody>
</table>

Participants were recruited through Japan Overseas Educational Services (JOES). They were enrolled in an English maintenance course offered by JOES. The course took place every Saturday for 90 minutes in central Tokyo. Native English speakers taught the course and its aim was to maintain the children’s English ability after returning to Japan. The Bilingual Language Experience Calculator (BiLEC) (Unsworth, 2016) was administered to the parents twice (first and second rounds of testing) to quantify the children’s language exposure when they were in a foreign country and a year after their return to Japan. This
questionnaire has been widely used in previous studies to measure experiential variables related to bilingualism (Kan & Schmid, 2019; Persici, Vihman, Burro, & Majorano, 2019; Stoehr, Van Hell, & Fikkert, 2019). Moreover, relative to other questionnaires that use Likert scales to measure exposure, BiLEC offers the advantage of quantifying exposure through specific algorithms, allowing for precise measures of exposure by considering various background information, rather than using simple ordinal rankings. Table 2 illustrates the mean language exposure to English and Japanese for home only and home, school, and extra activities (labelled ‘Extra’) for the first round and second round of testing. Children were mainly exposed to Japanese at home, regardless of whether they lived in a foreign country (Round 1: 85.6%) or in Japan (Round 2: 96.8%). English exposure (i.e., Extra) decreased dramatically from 46.8% exposure rate to 4.5% by the second round of testing. In contrast, Japanese exposure increased from 53.2% to 95.5%.

Table 2: Summary of BiLEC variables split by Language
‘Round 1’ indicates language exposures of when the children lived abroad and ‘Round 2’ indicates exposures of when the children returned to Japan; ‘Home’ indicates exposure at home and ‘Extra’ indicates exposures at home, school, and extra activities; the numbers are all in percentage

<table>
<thead>
<tr>
<th>Language (L2)</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 1</th>
<th>Round 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Home</td>
<td>Home</td>
<td>Extra</td>
<td>Extra</td>
</tr>
<tr>
<td>English</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>14.4</td>
<td>3.2</td>
<td>46.8</td>
<td>4.5</td>
</tr>
<tr>
<td>SD</td>
<td>18.2</td>
<td>7.9</td>
<td>12.1</td>
<td>5.6</td>
</tr>
<tr>
<td>Japanese</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>85.6</td>
<td>96.8</td>
<td>53.2</td>
<td>95.5</td>
</tr>
<tr>
<td>SD</td>
<td>19.9</td>
<td>12.6</td>
<td>10.8</td>
<td>8.5</td>
</tr>
</tbody>
</table>

2.2 Instruments and procedures

2.2.1 Language-switching

The language-switching task was programmed using E-prime (Psychology Software Tools, Pittsburgh, PA). It included a total of 84 trials in three blocks: two single language blocks (English and Japanese) and a mixed language block. In each trial, children had to...
name a picture presented on a computer screen in the language indicated by a visual cue. Participants were instructed to name the picture in English if they saw a British flag besides the target picture and in Japanese if they saw a Japanese flag with the target picture. A total of 50 black line drawn pictures (40 for single language blocks and 10 for mixed language blocks) were used in the study (Cycowicz, Friedman, Rothstein, & Snodgrass, 1997). We used pictures that elicited more than 90% accuracy rate in the name agreement data of six to seven year old L1 English children. All pictures were optically scanned, edited, and presented as black-on-white line drawings.

A trial began by presenting a white screen with a central black fixation cross (500ms), and then the stimulus picture (bottom of the screen) with the target language flag (on top of the screen). Given the fact that some studies found asymmetrical switch costs for shorter response-to-stimulus interval (i.e., time elapsed between the response and the presentation of the next stimulus; Mosca & Clahsen, 2016; Phillip et al., 2007), the researcher immediately pressed the space button to proceed to the next trial when the children had given a response. This measure was taken to minimize the response-to-stimulus interval and increase its sensitivity to observing asymmetrical switch costs. The same investigator administered the task for all participants in both rounds of testing to minimize administrator bias in determining the end of the target response. Any trials in which the response-to-stimulus interval was more than 1500ms (including the 500ms for presentation of the fixation cross) were excluded from the analysis. In addition, any trials where the button was pressed before the child finished naming the target stimulus were excluded. According to these exclusion criteria, 2.6% (80 trials) of the data were omitted from further analysis.

The order of the single language blocks (Japanese-English or English-Japanese) were counterbalanced, but always preceded the mixed language block. Each single language block
consisted of 20 test trials preceded by 5 practice trials. 20 different pictures were used in each single language block and the order of presentation was randomised. The mixed language block included 44 test trials divided into two sessions with 22 trials each, preceded by 16 practice trials. Eight target stimuli pictures were used in the mixed language block. That is, the same pictures were presented multiple times in a random order to the participants in the mixed language block. The mixed language block included both repetition trials (in which the language cue was the same as in the previous trial) and switch trials (in which the language cue differed from the previous trial). One-third of the trials were switch trials and the other two-third were repetition trials. There were equal numbers of repetition and switch trials for both languages. The language cue alternated on either every first, second, third, or fourth trial. The sequence of language cue was the same for every participant.

The participants were fitted with a microphone and a bilateral earphone set and seated in front of a computer screen. The instructions were given in the child’s most comfortable language in either Japanese or English by a Japanese-English bilingual researcher. The participants were reminded before each block to name the picture as quickly as possible. There was a short break between each single language block and mixed language block. In the single block, participants were not aware of the pictures that appeared in the experimental trial (the practice trial involved naming different items from the experimental trial). However, in the practice trial for the mixed block, the participants were asked to name the eight pictures that appeared in the experimental trial in both English and Japanese. Participants proceeded to the mixed language block only after they correctly named the eight pictures in English and Japanese in the practice trial. It took around 15 minutes for children to complete the picture-naming task. Built-in voice key functions were not used to extract reaction time, as they are easily triggered by background noises and other interfering sounds (e.g., err, umm) prior to
giving the actual response. Instead, naming latencies were measured by manually
determining the onset of the word through Checkvocal, a program that assists with processing
naming tasks data (Protopapas, 2007).

2.2.2 Simon task

The Simon task was constructed and administered using E-Prime (Psychology
Software Tools, Pittsburgh, PA). The Simon task was administered to measure children’s
ability to control interference and inhibit unwanted response. Participants were presented
with a white-background screen with either a frog or shoe on the two bottom corners of the
screen. The positions of the frog and the shoe were counterbalanced across participants. In
each trial, a target (either frog or shoe) was presented at the top of the screen, either on the
left or on the right. Participants were instructed to press the key that was located on the side
of the response picture that matched the target. In congruent trials, the position of the target
matched the position of the correct response picture (e.g., bottom left side = shoe, upper left
side = shoe). In incongruent trials, the target’s position and that of the correct response
picture did not match (e.g., bottom left side = shoe, upper right side = shoe). Thus,
incongruent trials required inhibiting the location of the target in order to respond based on its
identity.

Participants first completed 13 practice trials followed by a small break and the 40 test
trials. One-third of the test trials were congruent and two-thirds incongruent. Targets
disappeared after a certain amount of time, which was tailored to each participant’s own
response times in the practice trials, hence ensuring the task was equally challenging to all
participants regardless of individual differences in processing and motor speeds. The limit
was calculated by multiplying the mean reaction time in practice trials by 1.5. Instructions
were conveyed in the language that the child was most comfortable with. The children were instructed to push the correct keyboard button as quickly as possible. It took around 5 minutes for children to complete the task. The Simon effect was calculated by subtracting the average reaction time on the congruent trials from the incongruent trials.

2.3 Procedure

The language-switching task was always administered prior to the Simon task. The participants were all seen individually in a quiet room at the participants’ home by the investigator. The first round of testing took place in Summer 2016 and the second round of testing was administered a year later in Summer 2017. The exact same procedures were taken for the second round of testing. Since the current study was part of a larger project testing language attrition in bilingual children, other tasks not reported here were administered on the same day.

2.4 Data Analysis

In order to examine whether there was a significant change in RT for Simon and language-switching over time, we analysed the data using a linear mixed effect model available in the lme4 package (Baayen, Davidson, & Bates, 2008) in R (R Development Core Team, 2013). Before running the model, RTs were log-transformed to correct for normality. Inaccurate trials were also excluded from the dataset and RTs that were over or below three standard deviation from the mean were also omitted.

3. Results
3.1 Simon task

The mean RTs and accuracy of congruent and incongruent trials as well as the Simon effect in first and second round of testing are provided in Table 3. Simon task performance was analyzed with a model including Trial type (congruent or incongruent) and Time (first round or second round) with an interaction between Trial type and Time as fixed effects, and Subject and Item as random intercept. We applied sum coding for the predictors. The baseline intercept was set to congruent trials and first round of testing. Responses were significantly faster in congruent than incongruent trials (i.e., Simon effect), $E = .06, t = 5.82, p < .001$. There was also a decrease in RTs from first to second round of testing, $E = -.14, t = -12.15, p < .001$. As illustrated in Figure 1, RTs on both congruent and incongruent trials became faster over time. However, there was no interaction between Trial type and Time, $E = -.0007, t = -.04, p = .96$. That is, the Simon effect did not change from first to second round of testing.

Table 3: Descriptive statistics of the Simon task split by Time

<table>
<thead>
<tr>
<th></th>
<th>Congruent</th>
<th>Incongruent</th>
<th>Simon effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>First round</td>
<td>531 (144)</td>
<td>563 (129)</td>
<td>32 (15)</td>
</tr>
<tr>
<td></td>
<td>.91</td>
<td>.89</td>
<td></td>
</tr>
<tr>
<td>Second round</td>
<td>460 (125)</td>
<td>488 (110)</td>
<td>28 (15)</td>
</tr>
<tr>
<td></td>
<td>.91</td>
<td>.85</td>
<td></td>
</tr>
</tbody>
</table>

Note. Numbers without bracket indicate RT; numbers in brackets indicate SD; percentages below the RT and SD indicate accuracy
Figure 1: Interaction between Trial type and Time on the performance in the Simon task; error bars = standard error

3.2 Language-switching task

Mean reaction time, standard deviation, and accuracy for each Trial type (single, repetition, and switch) aggregated by Time (first round or second round) and Language (Japanese or English) are provided in Table 4.
Table 4: Descriptive statistics of each trial in the language-switching task split by Time and Language

<table>
<thead>
<tr>
<th>Language (L1)</th>
<th>Single</th>
<th>Repetition</th>
<th>Switch</th>
<th>Switching cost</th>
<th>Mixing cost</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Round</strong></td>
<td>1026 (537)</td>
<td>1166 (491)</td>
<td>1468 (736)</td>
<td>302 (245)</td>
<td>145 (45)</td>
</tr>
<tr>
<td></td>
<td>91.3</td>
<td>99.8</td>
<td>98.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Second Round</strong></td>
<td>931 (504)</td>
<td>1028 (414)</td>
<td>1131 (464)</td>
<td>103 (49)</td>
<td>96 (89)</td>
</tr>
<tr>
<td></td>
<td>91.3</td>
<td>99.8</td>
<td>98.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>English (L2)</strong></td>
<td>1186 (662)</td>
<td>1142 (541)</td>
<td>1365 (609)</td>
<td>224 (68)</td>
<td>-44 (122)</td>
</tr>
<tr>
<td></td>
<td>97.1</td>
<td>99.0</td>
<td>98.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>First Round</strong></td>
<td>1149 (657)</td>
<td>946 (396)</td>
<td>1096 (454)</td>
<td>151 (58)</td>
<td>-203 (262)</td>
</tr>
<tr>
<td></td>
<td>98</td>
<td>99.6</td>
<td>99.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Numbers without bracket indicate RT; numbers in brackets indicate SD; percentages below the RT and SD indicate accuracy.

The linear mixed effect model for language-switching included RT as a dependent variable; Language (Japanese or English), Trial type (single, repetition, and switch), and Time (first round or second round) as fixed effects, and Subject and Item as random intercept. We applied sum coding to the predictors. The baseline of the intercept was set to English, repetition trial, and first round of testing. The estimates of the model are provided in Table 5.
Responses became faster from first to second round of testing, \( E = -0.16, t = -7.48, p < .001 \). There was an interaction between Language and Trial type, \( E = -0.17, t = -4.10, p < .001 \) (Figure 2). Children’s naming latencies for Japanese words was significantly faster than for English words in the single blocks. However, in the mixed blocks (both repetition and switch) children were slower at naming items in Japanese than in English.

![Figure 2: Interaction between Language and Trial type on performance in the language-switching task; error bars = standard error](image.png)
There was a significant interaction between Time and Trial type, $E = .13$, $t = 4.48$, $p < .001$; that is, the magnitude of mixing costs increased from first to second round of testing. However, the significant interaction among Language, Time, and Trial type, $E = -.11$, $t = -2.75$, $p = .005$, suggests that the interaction between Time and Trial type was influenced by Language. Figure 3 illustrates the three-way interaction among Language, Time, and Trial type. The English mixing cost increased from first to second round of testing, but this effect was mainly due to the fact that the RTs of single trials in English did not decrease relative to the repetition trials. The mixing cost in Japanese did not change over time. The negative values for mixing cost in English (first round = -44 ms; second round = -203 ms) also showed that single trials were slower than repetition trials for both rounds of testing. This result was unexpected, given that single trials generally elicit faster responses than repetition trials as there is no need to switch between languages in the single block. In Japanese, however, the mixing cost values were positive (in the expected direction) and decreased over time (first round = 145 ms; second round = 96 ms).
3.3 Effect of bilingual experience and executive control on development of language control

Since a significant difference was only found in the magnitude of English mixing cost between the first to the second round in the language-switching task, we subsequently explored what factors could account for this difference. As shown in Figure 3, the change in English mixing cost appeared to be driven by single trials. The rate of RT decrease for the English single trials was not as steep as in the other trials. For instance, the RTs of Japanese single trials decreased over time by 95 ms, but English single decreased only by 37 ms.

In order to test the influence of bilingual experience and executive control in the observed changes in language control, we constructed two separate linear mixed effect
models, one with English single RT as a dependent variable, and the other with English repetition RT as a dependent variable. We included Time (first round and second round), English exposure, and Simon global RT change as fixed effects and Age at first round of testing as a covariate. Both models included Subject and Item as random intercept. We applied sum coding for the Time predictor and continuous predictors (English exposure and Simon global RT) were centered around the mean.

We used the English exposure measures at home, school, and extra activities at second round of testing for the English exposure variable. We took the measures from the second round of testing, as we expected that children who had more exposure in English when they were back in their L1 Japanese environment would be able to better enhance their performance on the English picture naming over time.

Simon global RT change was calculated by subtracting the global RT (RTs of both congruent and incongruent trials) of each participant from first to second round of testing. Higher values indicate faster performance on the Simon task over time. We used the global RT difference rather than the difference in Simon effect, as we found that Simon effect did not change over time and thus showed little variability. In contrast, there was a significant change in global RT, and most participants had faster RTs in the second round of testing (with the exception of two participants).

Before running the linear mixed effect model, we ran a Pearson product-moment correlations between the predictors (L2 exposure, Simon global RT change) and the covariate (Age at first round of testing) in order to ensure that there are no issues of multicollinearity. All correlations between these variables had low correlation coefficients but significant p-values: L2 Exposure and Simon global RT change, $r = .10, p <.001$; L2 Exposure and Age at
first round of testing, \( r = -.26, p < .001 \); Simon global RT change and Age at first round of testing, \( r = -.35, p < .001 \). Moreover, we used variance inflation factor (VIF) to confirm this, and all predictors had VIF values less than 5 (range 1.37 – 3.73), indicating that there is no issue of collinearity. The low correlation coefficient and VIF values indicate that these variables can be treated independently in further analyses.

We used backward elimination strategy to determine the most optimal model. We first constructed a full model with all predictors and interactions. We then eliminated interactions in a stepwise manner to test whether the new model without the specific interaction is significantly better than the old model with the interaction. We used a maximum likelihood ratio test to determine the more optimal model. We used backward elimination for the interactions only and the main effects were consistently kept in the model. We report the results of the optimal model for English single and English repetition trials in Table 6.

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Standard error</th>
<th>( t )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>6.95</td>
<td>.08</td>
<td>86.58</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Age</td>
<td>-.005</td>
<td>.001</td>
<td>-4.22</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>L2 exposure</td>
<td>-.04</td>
<td>.30</td>
<td>-1.36</td>
<td>.18</td>
</tr>
<tr>
<td>Round 2</td>
<td>.15</td>
<td>.04</td>
<td>3.20</td>
<td>.001**</td>
</tr>
</tbody>
</table>
Table 6: Estimated coefficients of the mixed effect model for English single and repetition trials

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Standard error</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simon</td>
<td>-.001</td>
<td>.0008</td>
<td>-1.37</td>
<td>.17</td>
</tr>
<tr>
<td>L2 exposure x Round2</td>
<td>-.60</td>
<td>.42</td>
<td>-3.74</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Simon x Round 2</td>
<td>-.001</td>
<td>.0005</td>
<td>-2.14</td>
<td>.03*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Variance</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>.02</td>
<td>.14</td>
</tr>
<tr>
<td>Item</td>
<td>.03</td>
<td>.18</td>
</tr>
</tbody>
</table>

**English repetition**

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Estimate</th>
<th>Standard error</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>6.85</td>
<td>.05</td>
<td>118.91</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Age</td>
<td>-.007</td>
<td>.001</td>
<td>-4.91</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>L2 exposure</td>
<td>.40</td>
<td>.29</td>
<td>1.00</td>
<td>.32</td>
</tr>
<tr>
<td>Simon</td>
<td>-.001</td>
<td>.0005</td>
<td>-1.32</td>
<td>.19</td>
</tr>
<tr>
<td>Round 2</td>
<td>-.07</td>
<td>.03</td>
<td>-2.40</td>
<td>.01*</td>
</tr>
<tr>
<td>Simon x Round 2</td>
<td>-.003</td>
<td>.0004</td>
<td>-3.05</td>
<td>.002**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Variance</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>.03</td>
<td>.17</td>
</tr>
<tr>
<td>Item</td>
<td>.002</td>
<td>.05</td>
</tr>
</tbody>
</table>

Table 6: Estimated coefficients of the mixed effect model for English single and repetition trials

*Note.* *p* < .05; **p < .01; ***p < .001

The output of the optimal model for English single trials showed that, Round of testing, L2 English exposure, and Simon global RT change predicted English single RT, as we found a significant interaction between L2 exposure and Round 2, \( E = -.60, t = -3.74, p < .001 \); as well as Simon and Round 2, \( E = -.001, t = -2.14, p = .03 \). The more English exposure the children had when they were back in Japan, the faster they had become at naming pictures in English in the single block (Figure 4). Similarly, bilinguals who became faster at responding in the Simon task also became faster at naming English items in the single block (Figure 5).
Figure 4: Interaction between English exposure and Time on English single trial performance

Figure 5: Interaction between Simon global RT difference and Time on English single trial performance
For English repetition trials, the only significant interaction was between Simon global RT change and Time, $E = -0.003 \, t = -3.05, p = .002$ (Figure 6), which was mainly due to the slope in the first round of testing. The larger the Simon global RT change was, the slower the RTs were in the first round of testing. In other words, children whose Simon performance increased the most responded faster in the English repetition over time, but this was mainly due to the fact that they started off with slower RTs in the first round of testing.

![Figure 6: Interaction between Simon global RT difference and Time on English repetition trial performances](image)

4. Discussion

This study investigated whether bilingual experience and/or executive control predict the development of bilingual returnee children’s language control. The findings show that magnitude of English mixing cost increased over time. By analyzing the English single trials
and the English mixed repetition trials separately, we found that both L2 exposure and
development in executive control modulated the change in English single trials, whereas
development in executive control (but not L2 exposure) predicted the change in English
mixed repetition trials.

We will first discuss the results of the Simon task, used to measure executive control.
Surprisingly, the magnitude of the Simon effect did not change over time. However, the
global RT (RTs of congruent and incongruent trials) significantly decreased from first to
second round of testing. The time span of one year between first and second round of testing
may have been too short to reveal a significant change in inhibition. Alternatively, inhibitory
skills may have already reached adult-like level for the bilinguals in the current study since
they were older children with a mean age of 9;8 (in the first round). Consistently, greatest
progress in inhibition (measured through frontal lobe functioning) is usually observed before
eight years of age (Best & Miller, 2010; Romine & Reynolds, 2005).

The lack of change for the Simon effect was due to similarly steep decline in RTs for
both congruent and incongruent trials over time. Consistently, a prior study reported faster
global RT in bilingual than monolingual children, but no difference in Simon effect (Martin-
Rhee & Bialystok, 2008). Although measures of inhibition in the Simon task are typically
defined as the RT difference between congruent and incongruent trials, global RT may reflect
the initial ability to control attention to complex stimuli (Martin-Rhee & Bialystok, 2008).
Specifically, bilingual children’s advantage in global RT may reflect greater ability to
monitor two sets of competing stimuli (Bialystok, 2006; Bialystok, Craik, Klein, &
Viswanathan, 2004; Costa, Hernandez, & Sebastian-Galles, 2008; Hichley & Klein, 2011;
Poarch & van Hell, 2012) or greater ability to handle tasks or trials of different types (Costa,
Alternatively, global RT in non-linguistic interference tasks may reflect processing costs; that is, a measure that is related to the speed, quantity, and quality of output in a mental task (Anderson, 2002). Indeed, processing speed may be an underlying construct that is reflected in all dimensions of the executive control and is not necessarily regarded as the “classic” executive control mechanism such as inhibition, switching, or updating (Salthouse, 1996, 2005). For instance, the global RT advantage found in Bialystok et al. (2004) has been interpreted as an indicator of enhanced processing speed (Diamond, 2013). Following this, a ‘neutral’ condition, where there is no need to monitor and control conflicts, could have helped differentiate the cognitive mechanisms behind global RT (Paap & Greenberg, 2013). Unfortunately, the lack of such a neutral condition in our study prevents making definite conclusions as to whether global RT decrease between the two testing sessions reflects improvement in processing cost or monitoring. Future studies should incorporate other tasks, such as Dimensional Change Card Sort (DCCS), which includes a neutral condition where participants must sort stimuli exclusively by color or shape—similar to the paradigm used in the language-switching task. However, the fact that we found a positive relationship between improvement in global RT and faster response in the English single block—which require no conflict monitoring between two languages—points towards the interpretation that processing cost is more likely to be manifested in global RT.

The findings from the language-switching task showed a significant interaction between Trial type and Language. In other words, while the bilinguals performed faster in Japanese than in English for the single block, they were slower in the mixed block. L1 global slowing effect was present in both time points, despite no differences in L1 and L2 switch costs across time. Similar L1 global slowing effects have been found in other language-switching studies (Costa & Santesteban, 2004; Costa et al., 2006; Gollan & Ferreira, 2009;
Misra, Guo, Bobb, & Kroll, 2012). In these studies, bilinguals were slower at L1 naming than L2 naming during language switching (but not when they did not have to switch between languages). L1 global slowing effect may reflect sustained language suppression, as suggested by ERP findings showing that L1 inhibition is applied not only in the immediate switch to the L2 but also persists in later L2 production (Misra et al., 2012). Our results lend support to the claim that the L1 global slowing effect emerges only when the bilinguals display symmetric switch costs (Costa et al., 2006).

Most importantly, English mixing cost increased over time, and this effect derives from the lower rate of decline in English single trials compared to repetition trials. Specifically, while there was on average -196ms difference from first to second round of testing for the English repetition, RTs on English single trials decreased only by 37 ms. Recall that in the single block trials, the bilinguals had to name 20 different pictures in each language. This involves language-specific process of matching 20 different concepts to their appropriate forms. Since the efficiency of mapping concept to form is determined by the use of the target language, (“weaker links” hypothesis: Gollan & Acenas, 2004; Gollan, Montoya, & Bonanni, 2005; Gollan & Silverberg, 2001), children who continued to receive English exposure in Japan may have been able to maintain the link between concept and form, and therefore experience less effect of attrition in English naming. In contrast, in the mixed block trials, the bilinguals had to name eight pictures that appeared repeatedly. Furthermore, they could only proceed to the experimental trial if they named all eight pictures correctly in the practice trial (this was not the case for the single block). In other words, we made sure that the link between concept and form was already established in the mixed block. This difference in experimental setup may explain why English exposure/use influenced the degree of change in the English single but not in the repetition trials. Taken together, our
findings suggest that the amount of L2 exposure influenced the performance on a specific task that require greater language-specific processing.

In addition to the role of language exposure, we were interested in how executive control predicts the development of language control. Our results showed that measures of executive control (i.e., difference in global RT on the Simon task) predicted the rate of development in both English single and English repetition trials. That is, bilingual children who increased their performance on the Simon task also became faster at English naming for both types of trials over time. Nevertheless, it should be noted that the bilinguals who experienced greater development in the Simon task started off (in first round of testing) with slower RTs in the English repetition trials. Therefore, their greater development in the English repetition trials may be motivated by the fact that they had more ‘room for improvement’, compared to others who were already fast at responding in the first round of testing.

In sum, our study shows that the change in English mixing cost over time can be explained by dual effects of bilingual experience and executive control. Bilingual experience especially influenced the change of mixing cost and in particular single block trials (but not mixed block) which involved naming many novel pictures, calling for language-specific processes that are mediated by language use. As the returnees experienced a dramatic decrease in their English exposure, this yielded only a moderate decline in RT for the English single trials, in comparison to a steep decline in English repetition trials; resulting in a larger mixing cost over time. However, it would be interesting to see whether this pattern would hold if the mixed blocks included as many novel pictures as the single blocks. In such a case, we would expect language exposure to also influence the performance in the repetition as well as switch trials. Furthermore, it is of interest to examine the relationship between
language control and executive control in other types of bilinguals who receive equal amounts of input in the two languages or are exposed to two languages from birth, in order to test the generalizability of our findings in populations of bilingual children other than returnees.

Most importantly, general cognitive development—whether it is monitoring or processing cost—predicted the performance of both English single as well as English repetition trials, suggesting partial overlap between executive control and language control. Unlike previous studies that looked at age-related changes of language control and executive control by comparing different age groups (Calabria et al., 2015; Weissberger et al., 2012), our longitudinal study shows that there is indeed a relationship between the developmental trajectories of language control and executive control. Even a short time span of one year was enough to reveal an interaction between the two control mechanisms. This is consistent with prior evidence for a direct correlation between language control and executive control in bilingual children (ages 5–7; Gross & Kaushanskaya, 2018). In that study, overall naming speed in the language-switching task was associated with shifting skills in DCCS; however, this relationship was only apparent in the children’s non-dominant language. In our study, we also found that the mixing costs of the children’s non-dominant language (L2 English) changed over time and were modulated by their improvement in the non-linguistic interference task. Stronger relationship between language control and executive control were observed for the non-dominant language, probably because use of the weaker language necessitates inhibiting the dominant language to a greater extent than use of the dominant language inhibiting the non-dominant one.

A hot debate in the current bilingual literature focuses on how language influences general cognition. Our study contributes to the literature by demonstrating that the
directionality of the relationship can be also explained the other way around—cognition can modulate language in the course of the children’s development. This sheds light on a fundamental question in bilingualism and cognition: does having two languages yield general cognitive benefits or does having advanced cognitive skills make one a better bilingual? Although our findings speak to the latter, this question remains open and further research is needed to investigate the relationship between these two key skill sets in a longitudinal manner and over a longer period of time.

5. Conclusion

In conclusion, although recent research has focused on how language control contributes to executive control, our study shows that this relationship is not unidirectional—in fact, executive control also has predictive power over language control in bilingual development. Our findings highlight the importance of considering different aspects of the bilingual experience when examining the development of language control abilities in children, including language-specific processes and domain-general processes involved in language control. Such an approach is especially promising to further our understanding of how bilingual experiences and executive control shape the children’s abilities to acquire, use, and maintain two languages.
References


