Bilingualism: Language and Cognition 22 (2), 2019, 366–382 © Cambridge University Press 2018 doi:10.1017/S1366728918000299

Code-switching does not predict Executive Function performance in proficient bilingual children: Bilingualism does*

CARISSA KANG
Cornell University
BARBARA LUST
Cornell University

(Received: February 14, 2017; final revision received: January 19, 2018; accepted: January 24, 2018; first published online 22 April 2018)

Previous studies of bilingual adults have suggested that bilinguals' experience with code-switching (CS) contributes to superior executive function (EF) abilities. We tested a highly bilingual developing population in Singapore, a multilingual country where CS occurs pervasively. We obtained CS and EF measures from 43 English–Chinese 8-year-old children (27 females, M = 100 months). We measured spontaneous CS with a novel task and EF in terms of task-switching (Semantic Fluency) and inhibitory control (Stroop task in both languages). Contrary to previous work, CS performance did not significantly predict EF performance in either case. Rather, bilingual language proficiency, i.e., degree of bilingualism (as measured by direct proficiency tests and parents' estimates of daily language use and exposure of both languages) influenced EF performance. Accordingly, the relationship between CS and EF may be more indirect and non-necessary than previously assumed.

Keywords: bilingualism, executive function, code-switching, bilingual language proficiency

1. Introduction

Code-switching (CS) is a natural and central feature of bilinguals' everyday lives. As Ansaldo (2010) has acknowledged, "multilingual contexts in which different languages are negotiated on a daily basis, and where language contact and contact languages are ubiquitous, are in fact quite common in human history." (p.615). Although there are varying definitions of CS (e.g., Muysken, 2000), we will define CS here as the spontaneous switching between two languages, or mixing of elements from both languages within a single speech event (Brice & Anderson, 1999; Nicoladis & Genesee 1997; Yow, Patrycia & Flynn, 2016; Yow, Tan & Flynn, 2017). CS is a complex phenomenon, involving cognitive and social as well as linguistic components (Kootstra, 2015). It requires precise modulation of more than one linguistic system (Meisel, 1994; Perez-Leroux, O'Rourke & Gretchen, 2014) and at the same time consideration of social "interactional context" (Green & Li, 2014; cf. Ansaldo, 2010). Different forms of interactive CS may reflect different types of cognitive and social control (Green & Abutalebi, 2013; Green & Li, 2014). Yet, CS appears effortless in many bilinguals.

Given that both of a bilingual's languages are always activated during listening, reading and production (e.g., Marian & Spivey, 2003; Kroll, Bobb & Wodniecka, 2006), CS becomes a natural possibility. However, processing costs have been found when adult bilinguals switch between two languages in lab-based language switching paradigms (e.g., Costa & Santesteban, 2004; Meuter & Allport, 1999). This discrepancy is important because language switching costs have been hypothesized to underlie a relation between bilingualism and EF, potentially explaining reported superior executive function (EF) in bilinguals (Abutalebi & Green, 2008; Barac & Bialystok, 2012).

In lab experiments, bilinguals are often required to name digits or pictures in their first (L1) or second (L2) language depending on external cues (e.g., color). A switch cost is calculated as the difference in naming latencies between switch trials (switching from L1 to L2 or vice versa) and non-switch trials. Asymmetrical switch costs have been reported where smaller switch costs were

* We thank our undergraduate research assistants at Cornell University – Cynthia Li, Yingxuan Chen and Jing Zheng – for help with reliability checking and transcribing, as well as the parents, school and participants who took part in this study. We also thank Prof. Felix Thoemmes for his consultation on statistical analyses. Finally, we would like to thank the two reviewers for their insightful comments on the paper. This research was supported by grants from the Cornell University Mario Einaudi Center for International Studies, Cornell University Human Development dissertation grant and Cornell University Cognitive Science research grant.

Address for correspondence:
Carissa Kang, Department of Human Development, Martha Van Rensselaer Hall, Ithaca, NY, 14853
ck577@cornell.edu
Supplementary material can be found online at https://doi.org/10.1017/S1366728918000299

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

observed when bilinguals switched from their dominant L1 to their less proficient L2 than the other way around (L2 to L1) (Meuter & Allport, 1999). It has been argued that naming in the weaker L2 necessitates a stronger inhibition of the more dominant L1 competitor, and the inhibition persists into the next switch trial, resulting in greater RT to name the digit in the L1 (Green, 1998; Meuter & Allport, 1999). These lab results appear intuitively paradoxical, given the pervasive and seemingly natural role of CS in bilingualism.

Recent research has begun to resolve this discrepancy by providing evidence that switching is not always costly for bilinguals. These accounts attribute the observed costs in previous studies to the nature of the task in lab-based language switching paradigms, where bilinguals are forced to switch as they are instructed on which language to use across each trial. They argue that these externally-generated switches prompt bilinguals to adopt inefficient and costly strategies not representative of language switching in daily lives. When switching is under the speaker's control, switch costs have been found to be reduced or absent (Gullberg, Indefrey & Muysken, 2009; Gollan & Ferreira, 2009; Kleinman & Gollan, 2016). Moreover, when sufficient sentence context is available, the context may provide information that reduces or eliminates processing costs (Gullifer, Kroll & Dussias, 2013).

Yim and Bialystok (2012) were one of the first who studied the relationship between conversational CS and one component of EF – task-switching. They examined whether greater CS frequency correlated with better task-switching performance in Cantonese–English bilingual undergraduates. Conversational CS was measured using a semi-structured task where participants were instructed to talk about various topics. To prime CS, the experimenter intentionally introduced a topic in both languages, and the content of topics was Chinese-themed (e.g., experiences during Chinese New Year) to encourage switching between both languages. Those who switched more during this conversational CS task incurred smaller costs during verbal task-switching (e.g., Semantic Fluency task).

To our knowledge, although there has been considerable work on the relationship between EF and bilingualism in children (Poarch & Van Hell, 2012; see Valian, 2015 for review), little previous work has specifically investigated the developmental relationship between CS and EF in children (See Yow & Li, 2015 for an exception). Examining the relationship between CS and EF in a developing population may provide critical insights into how the relationship may be foundationally related. In this paper, we investigate relations between CS and EF in a highly bilingual developing population – Singaporean English–Chinese bilingual children – where CS is prolific even in childhood. Here, we develop a new and more naturalistic methodology to measure CS.

To measure CS proficiency, we developed a CS task in a discourse format that allows participants to code-switch on their own accord, attempting to approximate internally-generated CS and to mirror the CS that occurs in bilinguals' everyday discourse. Our CS task comprises two conditions – inter- and intra-sentential. We included both these conditions since previous work has found that both types of switches may be associated with different forms of cognitive control processes (Green & Li, 2014). In our study, we define inter-sentential CS as when a speaker's utterance in one language is followed simultaneously by an utterance in another language (Genesee, Nicoladis & Paradis, 1995). The example below shows an inter-sentential CS from an English to Chinese utterance.

A: My mom has short hair and a very bad temper.
A: 我 最 喜欢和 妈妈 看 故事书
My most like with mother see storybook
(My favorite thing to do with my mother is to read storybooks.)

On the other hand, intra-sentential CS involves insertions of words or phrases from one language into the other language. In this simple example, the English noun "holiday" is inserted in a Chinese sentence.

A: 我 最 喜欢 的 holiday 是 新 年
My most like NOM is New Year
(My favorite holiday is New Year (NOM: nominative))

To explore how CS relates to EF, we examine how different components of CS relate to switch costs (e.g., verbal task-switching test) and inhibitory control (e.g., Stroop task, a common measure of verbal inhibitory control). Assuming a domain-specific relationship between CS and EF (i.e., both are verbal in nature), in this paper, we first focus only on the relationship between CS and verbal EF (see Kang, 2017 for the initial study of the relationship between CS and non-verbal EF). In addition, while previous studies included bilinguals with various proficiency types in their samples (e.g., a mix of Spanish- and English-dominant and balanced bilinguals (Kleinman & Gollan, 2016)), bilingual proficiency types may matter in assessing EF (e.g., Yow & Li, 2015). Thus, we restrict our population to proficient bilinguals. As advantages in EF have been primarily observed in balanced bilinguals (i.e., exposed to both languages and speak both well), we constrain and assess bilingual proficiency.

We assess bilingual proficiency through: i) a parental language background questionnaire (a subset of the Virtual Linguistics Lab Child Multilingualism Questionnaire (VLL CMQ, 2016) that was supplemented with an additional section on CS, ii) school exam scores in both languages, and iii) a brief production task (Elicited Imitation of complex sentences) in both languages (Blume & Lust, 2016; Lust, Flynn & Foley, 1996; Lust, Flynn,

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

Blume, Park, Kang & Kim, 2016). If engaging in CS does not tax EF when internally generated, we predict that in a population of bilinguals who code-switch prolifically, CS ability may not significantly correlate with EF performance. If so, then the relation between CS and EF may not be a direct one as sometimes suggested (e.g., Soveri, Rodriguez-Fornells & Laine, 2011; Prior & Gollan, 2011; Yim & Bialystok, 2012), but a more indirect and non-necessary one (Kleinman & Gollan, 2016).

2. Materials and method

2.1 Participants

Participants were recruited in Singapore, a country with a mandatory bilingual education policy. In Singapore, English is the official medium of instruction, and children are taught all subjects in English at school, except for their 'Mother Tongue' classes. The 'Mother Tongue' is usually assigned based on one's father's ethnicity (i.e., Mandarin for Chinese, Bahasa Melayu for Malays, etc.) and is taught as a secondary language at school. In homes, children are regularly exposed to both languages in most cases. In addition, Singlish is pervasive, a colloquial form of Singapore English, considered by some a 'creole', whose lexicon and grammar engage English, Chinese and Malay (Ansaldo, 2009; Leimgruber, 2011).

In this study, we recruit English–Chinese bilingual children, as the Chinese form the majority of the population (74.1%) (Department of Statistics, Ministry of Trade & Industry, Republic of Singapore, 2010). Pervasive and prolific CS has been reported for Singaporean English–Chinese adults (Xu, Chew & Chen, 1998) and English–Chinese Singaporean preschoolers (Yow et al., 2016). While CS occurs most frequently between Chinese and English in this population (Tan, 2004; Lee, 2005), it is also common to find CS involving other languages, dialects (e.g., Malay, Hokkien and Singlish).

Fifty-four 8-year-old English–Chinese bilingual Chinese children from the same primary school participated in this study. These were all Primary Two students (British system; almost equivalent to second grade of elementary school in U.S.) across seven different classes. All children were educated in English and Mandarin daily. As we needed to have the same language proficiency scores (e.g., school exam scores) for our study, children could only be recruited from one school. The principal distributed the consent forms to random students from each class in Primary Two, and we received approval from 54 parents. Out of 54 children, 11 children were involved in the pilot study to test our measures on this age group. The remaining 43 children (16 males; $M = 100.00$ months, $SD = 3.32$) participated in the actual study. Forty-two parents completed the VLL Child Multilingualism Questionnaire (CMQ; 2016) and confirmed that the sample spoke both English and Chinese at school, at home, and other places. All analyses were based on 43 children, except for analyses related to the VLL CMQ.

2.1.1 Language proficiency scores

In end-of-year examination scores for both English and Chinese languages from the school, each language had Oral and Written scores. English and Chinese Oral scores (16 and 30 points respectively; converted to 100%) included the following: i) Reading passage (e.g., student reads a prepared passage and is graded on vocabulary, pronunciation, articulation, fluency, rhythm and expressiveness), ii) Picture description, and iii) Conversation topic (e.g., teacher converses with student on a few topics; each student is graded on clarity of expression (fluency, language use, pronunciation) and engagement in conversation). For the Oral component, the average for English was 77.83% ($SD = 9.68$) and 82.54% for Chinese ($SD = 10.51$). Oral Chinese was significantly higher than Oral English, $t(41) = 2.41$, $p < .05$. As for the Written component, the English Written exam (40 points total, converted to 100%) comprised grammar, vocabulary, sentence synthesis and transformation, and an open-ended comprehension test. The average score for this was 90.00% ($SD = 5.68$). The Chinese Written exam (70 points total, converted to 100%) comprised grammar, vocabulary, sentence formation, comprehension (multiple choice and open-ended). The average score for this was 87.53 % ($SD = 7.51$). Written English scores were significantly higher than Written Chinese, $t(41) = 2.10$, $p < .05$.

We grouped bilinguals into (1) Balanced (high- and low-proficient) and (2) Unbalanced bilinguals (English- and Chinese-dominant) (selection criteria based on Rosselli, Ardila, Santisi, Arecco, Salvatierra, Conde & Lenis, 2002). Scores for the school language exam in English and Chinese were used as selection criteria. Participants with a difference between the Total English score (combined Oral and Written components) and Total Chinese score equal or lower than 11.50 (1 *SD* according to whole group scores) were represented as *Balanced Bilinguals* ($n = 34$), and those with difference between two exam scores greater than 11.50 points were *Unbalanced Bilinguals* ($n = 8$). We further split up the participants in the Balanced groups into high- and low-proficient balanced bilinguals. Bilinguals with a total score (English and Chinese) of 167 (group median) or higher were considered *high-proficient balanced bilinguals* ($n = 19$), while those with scores below group median were *low-proficient balanced bilinguals* ($n = 15$). Here, low-proficiency does not represent poor proficiency, but lower proficiency scores compared to the group median. The Unbalanced bilinguals were subdivided into English- or Chinese-dominant depending on their dominant language

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

performance for the school exam scores. *English-dominant unbalanced bilinguals* ($n = 6$) had higher total English scores, while *Chinese-dominant unbalanced bilinguals* ($n = 2$) had higher total Chinese scores.

2.1.2 Family background and VLL child multilingualism questionnaires

Parents also completed a family background questionnaire where they provided information about average monthly household income (1: Less than S$2,000, 2: S$2,001 – S$4,000, 3: S$4,001 –S$6,000, 4: S$6,000 – S$8,000, 5: S$8,001 – S$10,000, 6: More than S$10,000) and highest educational qualification for each parent (on a scale of 1 to 8 based on Singapore's education system, 1: Primary School Leaving Examination, 2: GCE 'N' Level, 3: GCE 'O' Level, 4: GCE 'A' Level/IB Diploma, 5: NITEC/Higher NITEC, 6: Polytechnic Diploma, 7: University Degree, 8: Graduate degree (post-university)). Participants ($n = 42$) came from families with an average monthly income of Level 4.96 (approx. S$8,001 – S$10,000) ($SD = 1.42$). Highest educational qualification (on a scale of 1 to 8) was 5.88 for mothers ($n = 42$; $SD = 1.92$) and 6.12 ($n = 41$; $SD = 1.89$) for fathers. This means that parents attained at least a Polytechnic diploma (one level below University degree). As the average median monthly household income in Singapore is S$8666 and approximately 42.9% of its population aged 25 and above has attained at least a diploma (Department of Statistics Singapore, 2015), this sample can be considered average to high SES in Singapore.

For language background (VLL Child Multilingualism Questionnaire, 2016), parents were asked to rate their child's overall language proficiency levels for English and Chinese on a scale of 1 to 7 (1: very poor, 2: poor, 3: fair, 4: functional, 5: good, 6: very good, 7: native-like for his/her age). All parents ($n = 41$) reported that their child was exposed to and used at least two languages on a daily basis, thus confirming that this sample is productively bilingual. Estimates of child's proficiency level for each language (out of 7: 7 being the highest), were significantly higher for English ($M = 5.38$, $SD = 0.80$) than Chinese ($M = 5.01$, $SD = 1.05$); $t(40) = 2.10$, $p < .05$. Parents were also asked to rate on a scale of 0 to 100 how bilingual/multilingual their child was (0: monolingual, 100: complete bilingual/multilingual). Average ratings ($n = 38$) of how bilingual their child was (0 to 100, 0: Monolingual, 100: complete bilingual or multilingual) was 72.11 ($SD = 24.95$).

Parents also reported estimates of their child's daily language exposure and use for their two languages. For each variable, the percentages add to 100% – that is, if a child is exposed to English 60%, Chinese exposure would be 40%. We computed two ratios of daily exposure to and daily use of English and Chinese (i.e., Daily Exposure/Use Ratio $= \frac{Percentage\ exposure\ or\ use\ [English]}{Percentage\ exposure\ or\ use\ [Chinese]}$. A score of 1 represents equal exposure/use to both languages, a score greater than 1 represents greater exposure/use to English and a score less than 1 represents greater exposure/use to Chinese). Estimates for average Daily Language Use Ratio was 6.17 ($SD = 15.38$) while average Daily Language Exposure Ratio was 2.85 ($SD = 3.21$). On a daily basis, most children were using English 6 times more than Chinese and exposed to English about 3 times more than Chinese.

For CS, parents were first asked to rate the following statement as accurately as possible, on a scale of 1 to 7 – "In general, your child changes from one language to another during conversation (either inserting words from another language, or alternating between two or more languages)" where 1: not true at all, 4: somewhat true, 7: very true. Parents were also asked to rate on a scale of 1 to 7 (1: Hardly ever, 4: Frequently, 7: All the time) how frequently their child changes from one language to another during conversation with primary caregiver, secondary caregiver, siblings (if any), and friends. Overall CS Frequency ($n = 40$) was 4.04 out of 7 ($SD = 1.63$) (i.e., "somewhat true" that child changes from one language to another during conversation). The averages for how frequently the child switches between languages with different people (on a scale of 1 to 7, 1: Hardly ever, 4: Frequently, 7: All the time) were: 3.21 out of 7 with primary caregiver ($n = 39$; $SD = 1.67$), 2.97 out of 7 with secondary caregiver ($n = 31$; $SD = 1.82$), 3.03 out of 7 with siblings ($n = 30$; $SD = 2.11$) and 3.06 out of 7 with friends ($n = 36$; $SD = 1.93$). Thus, this sample code-switches frequently across different situations.

2.2 Procedure

Parents completed the consent form and the Language and Family Background Questionnaires and returned these forms in sealed envelopes to the school. All participants were interviewed individually in a quiet room at the school. This was part of a larger study that took approximately 25–30 minutes to complete, where children completed an additional Executive Function task. Participants first completed the CS task, followed by 3 Executive Function tasks (Stroop, Semantic Fluency and a non-verbal EF task – Hearts and Flowers/Dots task (Davidson, Amso, Anderson & Diamond, 2006; reported in Kang, 2017); order counterbalanced). Finally, they completed the brief Elicited Imitation task assessing language production in both English and Chinese, with language order counterbalanced. Details for these tasks can be found in Section 2.3 below. We also obtained final exam scores for both English and Mandarin – both Written and Oral components from the school.

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

Participants were given a coloring set (worth $10) each for their participation.

2.3 Tasks.

2.3.1 Elicited Imitation (EI) task

To pursue converging evidence regarding children's language proficiency and to initiate direct testing of children's language competence (Lust et al., 2016) we used an Elicited Imitation (EI) task to assess a small set of various forms of complex sentences in English and Chinese, controlled for length and vocabulary across languages (see Appendix B for examples). This task has been used with experimental designs to assess grammatical knowledge in bilinguals and monolinguals (Gutierrez Ortiz, 2016; Lust et al., 1996). The small sample of sentences has been tested in full designs in larger studies of first and second language acquisition in both English and Chinese in both children and adults (Blume & Lust, 2016; Flynn, 1986; Lust & Chien, 1984).

We used an abridged version of this task from Flynn (1986) consisting of a battery of four stimuli sentences in each language to briefly assess knowledge of complex sentence formation using adverbial subordinate clauses with anaphora. Findings from previous research have shown that such direct production estimates may reveal significant differences in language proficiency over and above results from standardized tests (Gutierrez Ortiz, 2016; Flynn, 1986). These sentence types varied in embedding direction and anaphora type and direction. Participants were required to repeat each sentence after the experimenter under standardized conditions (Blume & Lust, 2016). After three practice trials, participants completed four experimental trials. Language order was counterbalanced. We computed percentage correct on 4 stimuli in each language, according to standardized scoring criteria for correctness (Blume & Lust, 2016). Major changes in syntax or semantics were considered incorrect.

2.3.2 Code-Switching (CS) task

In our novel CS task, there were two conditions – inter-sentential (inter-CS) and intra-sentential CS (intra-CS) – with eight questions in each condition (See Table 1 for examples). For both CS types, participants were told to answer the question as fast as possible and to say as much as they can. Importantly, they were NOT given any instructions regarding which language(s) to respond in. This ensured that they could choose the language they wanted to respond in (i.e., internally-generated response) for it to be more reflective of CS in natural discourse. In the Inter-CS condition, the experimenter asked questions in alternating languages (e.g., Question 1 in English, Question 2 in Chinese, Question 3 in English etc. as exemplified in Table 1). For this Inter-CS condition, there were two question orders – English- or Chinese-first. Order of language was counterbalanced. After each question, the child was given 20 seconds to respond. After 20 seconds, the experimenter interrupted the child (even if the child was still speaking) and asked another question in the other language, for example, "What were your favorite animals at the zoo? Why?" This condition allowed us to measure discourse-level CS, since the experimenter switches into another language; the child has to switch to the other language to comprehend the question and respond accordingly in the congruent language (or incongruent language, as they chose). In the Intra-sentential condition (Intra-CS), questions included cross-language insertions (exemplified in Table 1). There was only one order of questions in this condition.

We measured participants' DISCOURSE-LEVEL CS performance in the Inter-CS condition. That is, when the experimenter switched into another language, did the participant also do so and respond in the congruent language? We measured the proportion of trials that were responded to in the language congruent with the question. We also measured CS FLUENCY (RT), or the time taken for the child to respond to each question. This is in line with previous studies that also use RT as a CS measure in production tasks, when participants incur a RT cost as they switch between languages (see Abutalebi & Green, 2008 for a review). Here, we also use RT as the cost incurred to produce a congruent-language response to the experimenter's question, given that in the previous response (during the Inter-sentential CS condition), participant's response (and the question that was also asked) was in another language.

For Inter-CS Fluency, we did not include the RT for the response to the first question, since that did not involve any code-switching. Hence, there was a total of 7 trials of RT for Inter-CS Fluency. Finally, we measured CS INSERTION FREQUENCY – broadly defined by the number of times a child inserts words, phrases or other constituents from another language in a sentence. Given the design of our CS task, we expect that CS Insertion Frequency will be higher in the Intra- than the Inter-CS condition, where experimenter's questions model language insertions. Examples of different types of CS produced by children in the Intra-CS condition and how they were calculated can be found in Table 2 in Results.

2.3.3 Executive Function tasks: Semantic Fluency task

We tested Semantic Fluency, one of two tasks (the other being phonemic fluency) in the Verbal Fluency test that measures executive function and linguistic skills (Baldo & Shimamura, 1998; Monsell, 2003). We use the Semantic Fluency task as a measure of verbal task-switching as in Yim and Bialystok (2012). In this task, participants were given 30 seconds to list items that belong in a particular semantic category (e.g., types of

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

Table 1. *Examples of experimenter's questions in Inter- and Intra-CS. Italicized words are literal glosses.*

		Inter-sentential CS (Inter-CS)
Qn No.	Language	Questions
1	English	Have you been to the zoo? Who did you go with and what did you do at the zoo?
2	Chinese	在 动物园 里，最 喜欢 什么 动物? *In zoo inside most like what animal* 为什么 呢? *Why DP* (What was/were your favorite animal/s at the zoo? Why?)
3	English	Other than the zoo, what other places do you go with your family or on school excursions? Why do you like these places?
		Intra-sentential CS (Intra-CS)
1	Mixed	Can you tell me all about yourself? For example, 你 的 名字, 年龄, and hobbies? *you NOM name, age* (Can you tell me all about yourself? For example, your name, age, and hobbies?)
2	Mixed	可以告诉 我 你 最 喜欢 吃 什么 食物, *can tell me you most like eat what food* and why do you like to eat这 个 食物 呢 ? *this CLF food PRT* (Can you tell me what is your favorite food, and why you like to eat this food?)
3	Mixed	Can you describe你 的 妈妈? *you NOM mother* What does she look like, and 最 喜欢 和 *most like with* 妈妈 做 什么 呢? Why? *mother do what DP* (Can you describe your mother? What does she look like and what is your favorite thing to do with your mother?)

Key: Qn No: Question number, DP: discourse particle, NOM: nominative, CLF: classifier, PRT: particle.

sports). We reduced the Semantic Fluency task to 30 seconds instead of the usual 60 seconds since our pilot test revealed that it was challenging for children to come up with items for the entire 60 seconds. There are four conditions in this task – two blocked language and two mixed-language conditions. Four categories comprising animals, types of clothing, food and drinks, and body parts were tested. Participants were first asked to list items in each language (blocked language conditions). After this, they completed two mixed-language conditions, where items in each category had to be listed in alternate languages without repeating any item in both languages (e.g., for the category "Animals", an example of how an individual should respond is, "dog, 蛇 (snake), horse, 熊猫 (panda), giraffe, 鱼(fish), etc."). For the mixed-language conditions, participants were told which language to begin with, followed by the respective category. All instructions were provided in English (unless otherwise stated). None of the categories were repeated and these were counterbalanced for language and order.

The first two conditions were always the blocked-language conditions, followed by the two mixed-language conditions. For both blocked- and mixed-language conditions, starting language was counterbalanced – either English or Chinese first. Incorrect or inaccurate responses comprised responses were coded in the incorrect category, as were language and repetition errors (e.g., if the item was listed in the wrong language, or had already been repeated either in the same or the other language). A switch cost for each participant was based on the difference between the number of accurate responses on the blocked-language condition and the mixed-language conditions. This resembles calculating a global switch cost in a task-switching paradigm, but since there were no non-switch trials in the mixed-language conditions, this estimate may not be as precise.

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

Table 2. *Examples of children's CS insertions in the Intra-CS condition.*

Questions & Responses	CS Frequency
Question 可以 告诉 我 你 最 喜欢 吃 什么 食物, *can tell me you most like eat what food* and why do you like to eat 这 个 食物 呢? *this CLF food PRT* (Can you tell me what is your favorite food, and why do you like it?)	
Response A: I like to eat 板 面 from the 'kopitiam' because I find them very nice. *board noodles coffee shop* (I like eating "ban mian" (noodles) from the coffee shop)	2: 1 Chinese insertion, 1 Hokkien insertion ('kopitiam')
Question See Question No. 3 in Table 1, under "Intra-sentential CS".	
Response B: My mom has short hair and a very bad temper. 我 最 喜欢和 妈妈 看 故事书 *My most like with mother see storybook* (My most favorite thing to do is read storybooks with my mother.)	1: Switch from English to Chinese sentence
Question 你 的 爸爸 呢? What do you like doing with your 爸爸?为什么呢? *You NOM father PRT father why PRT* (What about) your father? What do you like doing with your father? Why?	
Response C: I like to play with my father too because sometime when my mother is busy she doesn't play with me at all. 我喜欢 和 我 的 爸爸 聊天" *I like with my NOM father chat* (I like chatting with my father.)	1: Switch from English to Chinese sentence
Question 学校recess 的 时候, 你 喜欢 和 你 的 朋友 play what games ? *School NOM time you like with you NOM friends* 可以告诉我 the rules of the game? 怎么 玩? *Can tell me how to play* (During recess time at school, what games do you like playing with your friends? Can you tell me the rules of the game? How do we play this game?)	
Response D: 我 和 我 的 朋友 喜欢 play uh chopsticks. 那 个 chopstick我们, *Me and me NOM friends like that CLF we* 我,是 那 个, like Addi(tion), like加 扣 分 *I is that CLF add deduct points* 那 个, 我们 要 uh tap 那 个finger。 那 个 ... *that CLF we have that CLF that CLF* (My friends and I like playing "chopsticks". That "chopsticks", we, I, is that, it's like Addition, like adding and deducting points. That (one), we have to tap that finger.)	5 English insertions: "play uh chopsticks" (Verb Phrase), "chopstick", "like Addition, like", "tap", "finger"

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

Table 2. *Continued.*

Questions & Responses	CS Frequency
Question	
你 最 喜欢 哪 一 个holiday呢?Christmas, 农历 新 年	
you most like which one CLF PRT Lunar New Year	
或者other holidays? 为什么 呢?	
or Why PRT	
(Which is your favorite holiday? Christmas, Lunar New Year or other holidays? Why?)	
Response E:	2 English insertions:
我 最 喜欢 的 holiday是Christmas.	"Holiday",
My most like NOM is	"Christmas"
(My favorite holiday is Christmas)	
Response F:	2 insertions: 1 Chinese
I like 新 年because I will receive a lot of ang pows	"新年", 1 Hokkien:
new year red packets	*"ang pows"*
(I like (Chinese) new year because I will receive a lot of red packets.)	

Key: DP: discourse particle, NOM: nominative, CLF: classifier, PRT: particle.

2.3.4 Executive Function tasks: Stroop task

We also administered a modified version of the Stroop task (Stroop, 1935), which has been used widely with monolinguals, bilinguals and trilinguals as a measure of verbal inhibitory control (Marian, Blumenfeld, Mizrahi, Kania & Cordes, 2013). In addition, Stroop has also been found to reflect bilingual proficiency – bilinguals with higher proficiency were found to be better at inhibiting interference compared to less proficient bilinguals (Zied, Philippe, Pinon, Havet-Thomassin, Ghislaine, Roy & Le Gall, 2004). In original versions of the Stroop, participants are shown color words in different ink colors and are asked to read aloud the color of the ink, ignoring the word. On congruent trials, the written word matches the color of the ink (e.g., blue written in blue ink). On incongruent trials, there is a mismatch between the ink color and the word (e.g., word "blue" written in black ink). Incongruent trials are typically more difficult, resulting in slower RT, demonstrating the "Stroop effect". This reflects competition between the automaticity in reading and the process of naming the color (MacLeod & MacDonald, 2000).

In our version of the Stroop, participants were shown 25 words (i.e., 5 words each in 5 rows) in different ink colors on a laptop screen, and asked to read aloud the color of the ink, ignoring the word, from left to right across each row from the top to the bottom. Participants completed one English and one Chinese Stroop task; with language order counterbalanced. There were 25 words in each language, with 7 colors in total, and 75% incongruent and 25% congruent trials in each block. Dependent measures were 1) Overall Accuracy (i.e., percentage of accurate trials over 25 words) and 2) Overall RT (i.e., total time taken to complete reading all 25 words) for each Stroop task. For accuracy, incorrect responses include incorrect colors read, as well as responses with self-corrections and hesitations. The critical dependent measure for the Stroop is the Stroop Interference Effect for Accuracy, or the difference in accuracy between the congruent vs. incongruent trials. Here, a smaller value represents less Stroop interference or better inhibitory control.

2.4 Data analysis

Two trained coders transcribed data from the CS and the Semantic Fluency task with reliability checked across coders. CS Fluency (RT data) was computed using Audacity. We evaluated the wavelength of the audio file and measured the time from the end of the experimenter's question (where the wavelength ends on the experimenter's last word) to the time of the participant's first word (i.e., start of the new wavelength for first word; see Appendix A for example). Non-words (e.g., "uh", "hmm", etc.) were not considered as first words. As this was a production task, we only excluded RT data for trials in which participants did not know what to say (e.g., they told the experimenter that they did not have anything to say, did not understand the question, or required many prompts for a response). Based on this, 20/301 trials (6.64%) from the Inter-CS and 13/344 trials (3.78%) from the Intra-CS condition were excluded. Two coders reviewed the CS Fluency data to ensure reliability and consistency of coding, and the final CS Fluency data were the average of both coders' RT data. The transcribed CS data are currently being entered in the Data Transcription and Analysis (DTA) tool (Blume, Flynn & Lust, 2012), a web application created for language research, collecting

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

both data and metadata in a structured database suitable for dissemination and ongoing collaborative analysis.

For CS Insertion Frequency, we counted the number of times of other-language insertions, without differentiating between different types of CS insertions (e.g., lexical, phrasal or clausal). Names (e.g., if a child was speaking in Chinese and inserted the name of an English teacher like, "Miss X"), communicators or interjections that were ambiguous – for example, those that can be used either in English or Chinese utterances such as "uh", "ah", "oh", or Singlish particles (e.g., "la", "meh"; see Rubdy, 2007) and onomatopoeia were excluded from the CS Insertion Frequency counts.

To assess if CS and bilingualism predict EF performance, we estimated three multiple linear regression models predicting EF performance for the Semantic Fluency and Stroop tasks (English and Chinese) from CS components and bilingual language proficiency scores. In addition, we performed correlational analyses to further examine how the individual CS and bilingual language proficiency scores correlate with EF performance. All statistical analyses were performed using SPSS.

3. Results

3.1 Elicited Imitation (EI) scores: Complex sentences

All children completed the full set of EI sentences in both English and Chinese. The average total score for English EI was 3.58 out of 4 ($SD = 0.63$) (89.53%) and for Chinese, 1.79 out of 4 ($SD = 0.94$) (44.77%). An example of an equivalent "Left-branching Backward Pronoun" (LB B Pron) in English and Chinese is below. In this example, the participant received a score of 1 for English and a score of 0 for Chinese because the child changed both the structure and lexicon for the sentence.

English: When he colored the books, Alex drank the milk.

Chinese: 当 他 给 书 涂 颜色 的
Dang1 ta1 gei3 shu1 tu2 yan2se de
时候, 哥哥 喝了 牛奶.
shi2hou4 gege he1 le niu2nai3
When he give book apply color NOM
time, elder brother drank milk.
(When he colored the book, elder brother drank milk.)

Response: 当 他 给 书 涂了 颜色
Dang1 ta1 gei3 shu1 tu2 le yan2se4
后, 当 哥哥 喝了 牛奶.
hou4 dang gege he1le4 niu2nai3
When he give book applied color after
when elder brother drank milk
(When he colored the book after, when brother drank milk.)

Thus, EI performance, like parents' estimates of daily language use and exposure, confirmed that the sample, although highly productively bilingual, were English-dominant to some degree. While standard academic proficiency scores and general behavior on the CS tasks (i.e., seamlessly switching back and forth between both languages throughout) confirm our sample to be proficient bilinguals, more precise estimates (both direct and indirect) begin to reveal a dominance of English over Chinese.

Nature of bilingualism

Based on exam scores and parents' ratings (e.g., proficiency level in both languages, level of bilingualism), this sample can be considered as productive bilinguals who are generally balanced but demonstrate an asymmetry in language control, generally favoring English – in terms of Written exams, Elicited Imitation, use and exposure estimates (although with dominance of Chinese for school Oral language exam scores).

3.2 Code-switching task

There were two orders in the Inter-CS condition – 21 participants completed order 1 (first question in English) while 22 completed order 2 (first question in Chinese). Analyses revealed that both groups did not differ significantly on their performance for all components of the CS task, all p's $> .10$. There were also no significant differences on CS performance across all bilingual groups (balanced and unbalanced). Please refer to Appendix C for the detailed breakdown of each CS component for each bilingual group (Balanced: low- and high-proficiency balanced, Unbalanced: combining both English- and Chinese-dominant bilinguals).

3.2.1. Discourse-level CS performance

In the Inter-CS condition, 93.90% of the trials (323/344) ($SD = 11.04$) were responded to in the congruent language. Almost all children switched to another language to match the experimenter when the experimenter switched the question in another language. Only one child did not switch language across all 8 questions in the inter-CS condition and responded to all questions in Chinese (note: in this case, experimenter asked the first question in Chinese). 6.10% (21/344) of incongruent language responses were produced by 34.88% (15/43) of the participants at least once. Of the 15 participants, 12 produced one incongruent language response. The remaining 3 produced 2, 3, and 4 incongruent responses out of 8 trials. English responses to Chinese questions occurred 52.38% of the time (11/21

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

trials), while the reverse occurred 47.62% of the time (10/21 trials).

3.2.2 CS Fluency

The average CS Fluency (average time taken to respond) for Inter- and Intra-CS was 1.47s ($SD = 0.77$) and 1.47s ($SD = 0.74$) respectively. A repeated-samples *t*-test revealed no significant differences in CS Fluency across CS conditions, $p > .05$. Within each CS condition (Inter- and Intra-CS), there were no significant differences in CS Fluency between the languages participants chose to respond in – Inter-CS English: $M = 1.42$s ($SD = 0.73$), Inter-CS Chinese: $M = 1.35$s ($SD = 0.77$), Intra-CS English: $M = 1.60$ ($SD = 1.17$), and Intra-CS Chinese: $M = 1.40$ ($SD = 0.60$); both *p*'s $> .05$. Participants responded just as quickly in English and Chinese. Regardless of the condition, the average time taken to respond was approximately similar and very fast (less than 1.5 seconds).

3.2.3 CS Insertion Frequency

In the Intra-CS condition, where the experimenter's question modeled CS insertions, 90.70% (39/43) of the participants code-switched by inserting words or phrases in their responses. In the Inter-CS condition, 37.21% (16/43) of the participants code-switched. The Intra-CS Insertion Frequency (i.e., number of times other-language words/phrases were inserted into responses) was 4.00 ($SD = 3.10$) while Inter-CS Insertion Frequency was 0.67 ($SD = 1.06$). Intra-CS Insertion Frequency was significantly higher than Inter-CS Insertion Frequency, $t(42) = 6.87$, $p < .001$.

Examples of the types of code-switches participants produced in the Intra-CS condition appear in Table 2. Most involved insertions of nouns (e.g., Responses E and F in Table 2), but there were also insertions of functional categories (e.g., connectives like 'but', 'and', 'then') as well as phrases (e.g., alternating between English and Chinese phrases as in Responses B and C in Table 2 and inserting of verb phrase; e.g., 'play chopsticks' in Response D).

In sum, for discourse-level CS performance (Inter-CS condition), a majority of the children switched to another language, and responded in the same language as the experimenter when the language of the question was switched. Participants were fast at responding across both CS conditions (less than 1.5 seconds for both), and responded as fast in both languages. Productive CS occurred in the Intra-CS condition, where 90.70% of the participants inserted other-language words or phrases in their responses. As expected, there was more productive CS Insertion Frequency in the Intra-CS condition.

Table 3. *Average number of correct responses in the blocked- and mixed-language conditions of the Semantic Fluency Task.*

Trial Type	Mean [95% CI]
Blocked-language condition	9.00 [8.50, 9.50]
English items	10.35 [9.32, 11.38]
Chinese items	7.65 [6.73, 8.57]
Mixed-language condition	5.87 [5.50, 6.25]
English first	5.98 [5.38, 6.58]
English items	3.23 [2.93, 3.53]
Chinese items	2.74 [2.40, 3.08]
Chinese first	5.77 [5.32, 6.22]
English items	2.72 [2.51, 2.93]
Chinese items	3.05 [2.77, 3.33]

3.3 Executive Function tasks

3.3.1 Semantic Fluency task

On the Semantic Fluency task, participants produced more items in the blocked-language conditions ($M = 9.00$) than the mixed-language conditions ($M = 5.87$), $t(42) = 10.78$, $p < .01$, cohering with previous literature (e.g., Yim & Bialystok, 2012). In the mixed-language conditions, repeated samples *t*-test revealed no significant differences in performance whether the participant started the block with English or Chinese, $p > .05$. The average switch cost was 3.13 ($SD = 1.90$) – approximately 3 fewer items produced on the mixed-language compared to the blocked language condition. Finally, in the blocked-language condition, a repeated samples *t*-test showed that participants performed significantly better on the English than Chinese block, $t(42) = 3.15$, $p < .01$ (see Table 3).

3.3.2 Stroop task

Participants performed better on the English compared to Chinese Stroop on all components of the Stroop task (see Table 4). English accuracy was higher (93.21%) than Chinese accuracy (90.51%). Overall, participants took less time to complete the English Stroop ($M = 38.03$s) than the Chinese Stroop ($M = 50.01$s). The Stroop Interference Effect (Accuracy) was smaller in the English (0.07) than Chinese Stroop (0.12); the difference in performance on congruent and incongruent trials was greater in the Chinese Stroop compared to English Stroop. English Overall Accuracy was significantly higher than Chinese, $t(42) = 2.06$, $p < .05$, while English Overall RT was significantly faster than Chinese, $t(42) = -5.72$, $p < .01$. Stroop Interference Effects were significantly smaller in the English than the Chinese Stroop, $t(42) = -2.11$, $p < .05$. Stroop performance reflects English dominance, since

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

Table 4. *Summary of components of English and Chinese Stroop task.*

Condition	English Mean (*SD*)	Chinese Mean (*SD*)
Overall Accuracy (%)	93.21 (8.92)	90.51 (8.51)
Overall RT (s)	38.03 (8.63)	50.01 (13.68)
Stroop Interference Effect (Acc)	0.07 (0.10)	0.12 (0.12)

Note: Overall Accuracy and RT refers to total accuracy and total time taken across all Stroop trials, while Stroop Interference (Acc: Accuracy) refers to difference in accuracy between congruent and incongruent trials.

children performed significantly better on all components of the English Stroop.

3.4 Correlational analyses

3.4.1 Relationship between CS and EF performance

We first conducted correlational analyses to examine how each CS component correlated with the various EF scores. The five CS components – discourse-level CS, Inter- and Intra CS Fluency and Intra-CS Insertion Frequency – were correlated with Semantic Fluency switch cost. Only Intra-CS Fluency was significantly correlated with Semantic Fluency switch cost, r = -0.33, $p < .05$. Those who responded faster on the Intra-CS condition incurred higher switch cost on Semantic Fluency. All other correlations between CS components and Semantic Fluency switch cost were insignificant, p's > .05. Correlational analysis between Intra-CS Insertion Frequency and Semantic Fluency switch cost revealed a negative relationship although this was not significant (r = -0.10), unlike in Yim and Bialystok (2012). When we analyzed correlational analyses between Intra-CS Insertion Frequency and switch cost across different groups of bilinguals (balanced and unbalanced bilinguals), results revealed relationships in opposite directions. The results for the correlations between Intra-CS Insertion Frequency and switch cost across different groups of bilinguals are as follows for each group: a) low-proficiency balanced (n = 15): r = -0.23, b) high-proficiency balanced (n = 19): r = -0.02, and c) unbalanced (English- and Chinese-dominant combined, n = 8): r = 0.24. Although these correlations were not significant and the sample sizes for each group were small, these analyses provide preliminary evidence for different CS-EF relationship depending on proficiency and level of bilingualism. Semantic Fluency switch cost and Intra-CS Insertion Frequency were the least correlated for the high-proficiency balanced bilinguals.

Regarding correlations between CS and Stroop performance, discourse-level CS did not significantly correlate with English and Chinese Stroop Interference, both p's > .10, although these correlations were negative (e.g., more congruent-language responses on Inter-CS, smaller Stroop Interference Effects). In addition, Inter- and Intra-CS Fluency and Frequency with English and Chinese Stroop Interference Effect (Accuracy) also were not significant, all p's > .10.

Lastly, correlational analyses between parents' estimates of children's CS Frequency with Semantic Fluency switch cost were also not significant p > .10. In general, our CS measures (based on our CS task and parental estimates) did not correlate significantly with verbal task-switching.

3.4.2 Relationship between bilingual language proficiency and EF performance

We next examined how bilingual language proficiency correlates with Semantic Fluency switch cost. School language proficiency scores (Average English and Chinese), daily language use and exposure ratios, as well as Elicited Imitation scores (in both languages) did not correlate significantly with Semantic Fluency switch cost, all p's > .05. In addition, CS components also did not significantly correlate with Stroop Interference Effect (Accuracy) in both languages.

On the other hand, factors related to language proficiency did correlate with Stroop Interference Effect in both languages. We correlated school language proficiency scores (average English and Chinese scores across Written and Oral components), Elicited Imitation scores (in both languages) as well as daily language use and exposure ratios with English and Chinese Stroop Interference Effects (Accuracy). Average English score (combined score for English Oral and Written) was significantly correlated with English Stroop Interference Effect, r = -0.40, $p < .01$. English Elicited Imitation was significantly correlated with English Stroop Interference Effect (Accuracy), r = -0.46, $p < .01$. Those who performed better on the English Elicited Imitation also had lower English Stroop Interference. Ratio of daily language use was also significantly correlated with Chinese Stroop Interference Effect (Accuracy), r = 0.39, $p < .05$. Those who used more English (or conversely, less Chinese) daily had higher Chinese Stroop Interference.

3.4.3. Relationship between CS and bilingual language proficiency

Next, we examined how bilingual language proficiency (as measured by school exam scores, Elicited Imitation and ratio of language use and exposure) correlates with CS performance (Discourse-level CS, Inter- and Intra-CS Fluency and Intra-CS Insertion Frequency). Oral Chinese was negatively correlated with Inter-CS Insertion

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

Frequency, $r = -0.35$, $p < .05$. Those with higher Chinese Oral scores produced fewer other-language insertions during Inter-CS (where no language insertions had been modeled). In addition, ratio of language use was also correlated with Inter-CS Insertion Frequency, $r = 0.69$, $p < .001$. Those whose parents reported greater use of English daily (compared to Chinese) also produced more other-language insertions during the Inter-CS condition.

Elicited Imitation scores were also significantly correlated with CS Fluency. English EI was positively correlated with Inter- and Intra-CS Fluency, $r = 0.33$ and 0.31 respectively (both p's $< .05$). That is, those who performed better on English EI took a longer time to respond to questions in both the Inter- and Intra-CS conditions. Finally, Chinese EI score was negatively correlated with Discourse-level CS, $r = -0.30$, $p < .05$. Those who performed better on the Chinese EI also produced less congruent-language responses.

3.4.4. Relationship between parents' estimates of child's CS frequency and CS measures

Next, we examined how parents' estimates of child's CS frequency (on a scale of 1 to 7) correlated with child's CS performance (Discourse-level CS, Inter- and Intra-CS Fluency and Intra-CS Insertion Frequency). None of these correlations were significant.

3.5. Regression Analyses: How do CS performance and bilingual academic proficiency together predict EF performance?

To examine if and how CS ability and bilingual academic proficiency predicted EF performance, we estimated multiple regression models predicting EF performance from CS and academic scores. As Written and Oral scores were significantly correlated within each language – English: $r = .30$ ($p = .05$) and Chinese: $r = .40$ ($p < .01$) – we combined the Written and Oral components for each language into Average English and Average Chinese scores.

We estimated three multiple regression models with different outcome variables for the EF tasks – Semantic Fluency switch cost (Model 1), English Stroop Interference Effect (Model 2) and Chinese Stroop Interference Effect (Model 3). In all three models, the predictors were the CS measures (Inter- and Intra-CS Fluency, and Intra-CS Insertion Frequency) and language proficiency (as defined by the school exam scores). For the CS predictors, we excluded Inter-CS Insertion Frequency, since only a small number of participants engaged in CS in the Inter-CS condition and the average Inter-CS Insertion Frequency was low ($M = 0.67$). The effects of CS performance and language proficiency on Semantic Fluency switch cost were insignificant, with an overall R^2 value of 0.24 ($SE = 1.82$, $F(6, 41) = 1.80$), $p > .10$ (Model 1). In this model, Average Chinese score was the only significant predictor of Semantic Fluency switch cost, $\beta = 0.09$ ($SE = 0.04$), $p < .05$; those with higher Chinese scores (Written and Oral combined) incurred higher switch cost on Semantic Fluency. None of the CS components significantly predicted Semantic Fluency switch cost, all p's $> .05$.

Overall, the effects of CS performance and language proficiency on English Stroop Interference were also insignificant, with an overall R^2 value of 0.19 ($SE = 0.10$, $F(6, 41) = 1.34$), $p > .10$ (Model 2). In this model, only Average English score was a significant predictor of English Stroop Interference, $\beta = -0.01$ ($SE = 0.003$), $p = .01$; those with higher Average English scores had lower English Stroop Interference. None of the CS components (Discourse-level CS, Average CS Fluency and Intra-CS Insertion Frequency) were significant predictors of English Stroop Interference, all p's $> .10$. Lastly, in Model 3, the effects of CS performance and language proficiency on Chinese Stroop Interference were also insignificant, with an overall R^2 value of 0.10 ($SE = 0.12$, $F(6, 41) = 0.68$) respectively, $p > .10$. In this model, none of the CS or school language proficiency scores were significant predictors, all p's $> .05$ (see Table 5).

In sum, all three regression models revealed that CS performance was not a significant predictor of EF performance – whether it was task-switching (Semantic Fluency) or inhibitory control (Stroop). However, various components of bilingual language proficiency (i.e., school language proficiency scores) did predict Semantic Fluency switch cost and English Stroop Interference Effect.

Both regression and correlational analyses confirmed that CS ability did not predict EF performance in task-switching and inhibitory control in our sample of proficient bilinguals. However, even though school language proficiency, elicited imitation and daily language use and exposure ratios in general did not correlate significantly with Semantic Fluency switch cost, those who performed better on overall English exam scores (Average English) and on English Elicited Imitation also displayed lower English Stroop Interference Effect. Those who used less Chinese on a daily basis also displayed higher Chinese Stroop Interference. Interestingly, whether children were English- or Chinese-dominant bilinguals influenced the direction of the relationship between CS Frequency and switch cost.

3.6 Summary of results

The Singaporean child population we studied was found to include proficient bilinguals and proficient

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

Table 5. *Summary of Multiple Regression Analyses for Variables predicting Semantic Fluency Switch Cost and Stroop Interference Effect Accuracy (both languages).*

Variable	Model 1 Semantic Fluency Switch Cost *B, SE*(*B*)	Model 2 Stroop Interference (English) *B, SE*(*B*)	Model 3 Stroop Interference (Chinese) *B, SE*(*B*)
Intercept	3.53 (4.98)	0.55 (0.27)	0.34 (0.34)
Discourse-level CS	0.32 (2.72)	0.04 (0.15)	−0.22 (0.19)
Inter-CS Fluency	−0.51 (0.59)	−0.02 (0.03)	−0.06 (0.04)
Intra-CS Fluency	−0.44 (0.84)	0.002 (0.04)	0.07 (0.06)
Intra-CS Insertion Frequency	−0.06 (0.09)	−0.000 (0.01)	0.002 (0.01)
Average English	−0.08 (0.05)	−0.01 (0.003)**	0.0031(0.003)
Average Chinese	0.09 (0.04)*	0.001 (0.002)	−0.002 (0.003)
R^2	0.24 ($SE = 1.82$)	0.19 ($SE = 0.10$)	0.10 ($SE = 0.12$)
F	1.80	1.34	0.68

*$p < .05$, **$p < .01$

code-switchers, as measured both by our indirect (i.e., parent reports) and direct measures (i.e., school language proficiency and elicited imitation scores). In our CS task, participants were fast and fluent at responding to questions across Inter- and Intra-CS conditions, whether they responded in English or Chinese. They were proficient at discourse-level CS performance, and could switch languages in less than 1.5 seconds whenever the experimenter switched. Critically, however, CS ability did not predict or correlate with EF performance. That is, better CS performance (in terms of higher proportion of congruent-language response for discourse-level CS), lower CS Fluency and higher CS Insertion Frequency did not correlate with lower Semantic Fluency switch cost nor with lower Stroop Interference Effects.

In contrast, analyses of the nature of bilingualism in this population through both direct and indirect measures did demonstrate effects of bilingualism on EF. Our population appeared to be seemingly balanced bilinguals, as reflected by the standardized exam scores and parents' ratings of language proficiency levels, as well as their flexible performance on all our tasks. On the other hand, careful analyses of parents' estimates of daily use and exposure of both languages as well as Elicited Imitation performance did suggest that our sample was English-dominant. Accordingly, performance on both Semantic Fluency and Stroop reflected English dominance – in terms of more items produced in the English condition compared to the Chinese condition on Semantic Fluency, and better performance on all components of the English Stroop. Chinese and English proficiency scores also significantly predicted Semantic Fluency switch cost and English Stroop Interference Effect respectively.

4. Discussion

The objective of this study was to investigate if a relation would hold between CS and EF in a developing population, specifically 8-year-old children still developing linguistically and cognitively. We examined whether there is a CS-EF relationship, especially in a population where CS occurs pervasively at both the individual and broader sociolinguistic level. Our results revealed that none of the CS components that we measured predicted EF performance in terms of the basic EF components of inhibitory control and task-switching. Our finding that CS ability does not predict inhibitory control (on the Stroop) accords with previous research revealing the parallel activation of both languages during language production and comprehension in adults (Dijkstra, 2005; Kroll et al., 2006). CS may not necessitate inhibition of one language or the other, especially when switching between languages appears to occur seamlessly in natural situations. These results also accord with a recent cognitive-social model of code-switching (Green & Li, 2014) wherein an "interactional context" displaying a "recurrent pattern of conversational exchanges within a community" (p. 502), such as we observe in Singapore, may not require suppression of competing languages, but a different mode of control which is more "open".

We also did not find a significant relationship between CS frequency and verbal switch cost. If, as hypothesized, language control processes change according to the continuing demands of the interactional context (e.g., requiring a 'cooperative' relation between languages in dense CS), then speakers should not incur any explicit costs during the production of language switching (Green & Abutalebi, 2013).

Although several previous studies have found that frequent language switching (either measured by

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

self-reports or other CS measures) correlates with smaller task switching costs (Prior & Gollan, 2011; Yim & Bialystok, 2012), we found only one anomalous correlation between Inter-CS Fluency (RT) and Semantic Fluency switch cost ($r = -0.33$). However, we do not know if those with lower switch cost i) took longer to respond (as we did not collect RT data for Semantic Fluency), ii) performed poorly across blocked and mixed conditions, or iii) performed almost just as well across blocked and mixed conditions. Thus, it is difficult to interpret this significant correlation.

Thus, CS may not recruit EF to the same extent in prolific code-switchers who are regularly being exposed to and engage in CS. This is a likely possibility in our population, given the high prevalence of CS across multiple languages in Singapore, including in children, resulting from the continual interactions of individuals across different ethnic groups. This was further confirmed in the language background questionnaires where parents reported that their children engaged in CS at home, in school and with friends. Interviews with the children further confirmed parents' reports. When asked why they engaged in CS so frequently, they explained it was "just a habit", or the meaning of some words were better captured in a particular language, or simply that they did not know or remember the word in the other language. Thus, the CS-EF relationship may differ depending on the broader sociolinguistic context; bilinguals from "different behavioral ecologies" may experience differences in control demands (Green, 1998; 2011). In addition, this hypothesis requires comparing non-habitual with habitual switchers. Given that our study only tested habitual switchers, more research is required to fully pursue this hypothesis.

The lack of a significant CS-EF relationship, which we have discovered, may also reflect, at least partially, the nature of our CS task. Gullifer et al. (2013) argued that inter-sentential CS (which is a type of CS that bilinguals engage in frequently) incurs no significant costs to lexical processing, compared to lexical switching across languages. Thus, the sentence context in our CS task might have further reduced processing costs for our participants, suggesting that the "experimental context" as well as the social context also plays a role in control responses.

In addition, self-generated or voluntary CS has been shown to be less costly and less taxing compared to involuntary CS. In our task, although it is the experimenter's language switching that triggers the speaker to code-switch, ultimately, the child decides whether to switch in response. In doing so, the child must consult the identity of the speaker (known to speak both languages fluently, just as the child does) and engage in a type of "identity alignment", argued to characterize "multilingual contexts in which different languages are negotiated on a daily basis" (Ansaldo, 2010, p.615), aligning identity and language. The high percentage of discourse-level CS found in our data attests to the natural and productive role of this process in our subjects, perhaps in keeping with the ethnically and linguistically heterogeneous cultural context of Singapore. This finding provides interesting information about the interactive alignment of CS, which, to our knowledge, has not yet been studied in children. The fact that children in the current study also show evidence of primed CS suggests that the mechanism of priming in bilingual language use is also at work in bilingual children.

Importantly, our results also revealed that in our multilingual multicultural population from Singapore, aspects of bilingual language proficiency do predict EF performance. Though we tested multiple measures of language proficiency, bilingual language proficiency, defined by academic proficiency (written and oral components), direct production tests and parental estimates of use and exposure were discovered to reveal previously unrecognized language asymmetries and such asymmetries proved to be a significant predictor of how bilingual children perform on EF tasks. The English dominance we discovered in our Singaporean sample was also found in a recent study in a highly proficient bilingual Singaporean sample of 5 to 7-year-old children (Yow et al., 2017). In recent work, intra-sentential code-switching has been found to be modulated by bilingual dominance as well as by grammatical structure and type of switching (Perez-Leroux et al., 2014). These results make clear that bilingual proficiency must be carefully taken into account in the assessment of the relationship between bilingualism and CS. They in fact raise the question of whether totally balanced bilingualism exists (Cutler, Mehler, Norris & Segui, 1992). They may help to explain some inconsistent findings in the field, where not all studies have found a bilingual EF advantage and issues have arisen as to what the underlying mechanisms are for the superior EF performance (e.g., Valian, 2015). Our results also provoke further research to more specifically differentiate between types of CS (Muysken's (2000) CS types), potentially differentiating between different cognitive control processes (e.g., Green & Li, 2014).

In conclusion, our results confirm that cognitive mechanisms related to code-switching may be flexible. The relationship between CS and EF may be more indirect and non-necessary than previously assumed. As has been speculated, in "dense CS contexts" speakers (and hearers) must "freely exploit the activation of both languages" and speakers may "have no need to avoid switching" (Green, 2011, p.2). At this time, we do not have a comprehensive theory of code-switching, which would be necessary for further development of this hypothesis. It is clear when one considers the social complexity of discourse mixing or the grammatical complexity of intra-sentential mixing that such a theory must be interdisciplinary: linking study of

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

cognitive-social mechanisms, grammatical mechanisms consulting language typology and structural and linear aspects of language grammars and constraints, basic language processing mechanisms and evaluation of the linguistic and cognitive competencies of the bilingual speaker-hearer, potentially supported by neuroscientific investigations (Abutalebi & Green, 2008).

In future research, our results motivate us to target the critical importance of the study of language in multilingual code-switching populations, such as Singapore. As has been suggested, "[it] may be the case that the natural state of human communication is to be found in multilingual, creative negotiations in which different codes are used simultaneously and with periodic shifts in dominant registers" (Ansaldo, 2010, p. 622). Here, the creative process of language acquisition may be more thoroughly revealed.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S1366728918000299

References

Abutalebi, J., & Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes*, *23*, 557–582.

Ansaldo, U. (2009). Contact Languages. Ecology and Evolution in Asia. Cambridge University Press.

Ansaldo, U. (2010). Identity alignment and language creation in multilingual communities. *Language Sciences*, *32*, 615–623.

Baldo, J. V., & Shimamura, A.P. (1998). Letter and category fluency in patients with frontal lobe lesions. *Neuropsychology*, *12*, 259–267.

Barac, R., & Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child Development*, *83*, 413–422.

Blume, M., Flynn, S., & Lust, B. (2012). Creating linked data for the interdisciplinary international collaborative study of language acquisition and use: Achievements and challenges of a new Virtual Linguistics Lab. In Christian Chiarcos, Sebastian Nordhoff, and Sebastian Hellmann (eds.) *Linked Data in Linguistics: Representing and Connecting Language Data and Language Metadata*, pp. 85—96. Heidelberg: Springer.

Blume, M., & Lust, B. (2016). Experimental Tasks for Generating Language Production Data. In *Research methods in language acquisition: Principles, procedures and practices* (pp. 119–136). Mouton de Gruyter and American Psychological Association, Washington D.C.

Brice, A., & Anderson, R. (1999) Code mixing in a young bilingual child. *Communication Disorders Quarterly*, *21*, 17–22.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*, 491–511.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, *24*, 381–410.

Davidson, M. C., Amso, A., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 year: Evidence from manipulations of memory, inhibition, and task-switching. *Neuropsychologia*, *44*, 2037–2078.

Department of Statistics, Ministry of Trade & Industry, Republic of Singapore. (2010). Census of population 2010 Statistical Release 1: Demographic characteristics, education, language and religion. Department of Statistics, Ministry of Trade and Industry. Retrieved from: https://www.singstat.gov.sg/docs/default-source/default-document-library/publications/publications_and_papers/cop2010/census_2010_release1/cop2010sr1.pdf

Department of Statistics Singapore (2015). Key Household Income Trends, 2015. [Retrieved from: https://www.singstat.gov.sg/docs/default-source/default-document-library/publications/publications_and_papers/household_income_and_expenditure/pp-s22.pdf]

Dijkstra, T. (2005). Bilingual word recognition and lexical access. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). New York: Oxford University Press.

Flynn, S. (1986). Production vs. comprehension: Differences in underlying competencies. *Studies in Second Language Acquisition*, *8*, 135–164.

Genesee, F., Nicoladis, E., & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language*, *22*, 611–631.

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 640–665.

Green, D. W. (1998). Mental control of the bilingual lexicosemantic system. *Bilingualism: Language and Cognition*, *1*, 67–81.

Green, D. W. (2011). Language control in different contexts: the behavioral ecology of bilingual speakers. *Frontiers in psychology*, *2*, 103.

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

Green, D. W., & Li, W. (2014). A control process model of code-switching. *Language, Cognition and Neuroscience*, *29*, 499–511.

Gullberg, M., Indefrey, P., & Muysken, P. (2009). Research techniques for the study of code-switching. In B. E. Bullock, & J. A. Toribio (Eds.), *The Cambridge handbook on linguistic code-switching* (pp. 21–39). Cambridge: Cambridge University Press.

Gullifer, J. W., Kroll, J. F., & Dussias, P. E. (2013). When language switching has no apparent cost: Lexical access in sentence context. *Frontiers in Psychology*, *4*, 1–13.

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

Gutierrez Ortiz, M.d.M. (2016). Assessing content and language integrated learning: A study of the second language acquisition of English in monolingual Andalusia (Spain). Unpublished Master's thesis, Cornell University, Ithaca, NY.

Kang, C. (2017). Exploration of code-switching as a mechanism underlying bilingual executive function performance: A study of 8-year-old English–Chinese Singaporean bilingual children. *Doctoral thesis*. Retrieved from ProQuest Dissertations and Theses database. (ID No. 10293).

Kleinman, D., & Gollan, T. H. (2016). Speaking two languages for the price of one: Bypassing language control mechanisms via accessibility-driven switches. *Psychological Science*, *27*, 700–714.

Kootstra, G. J. (2015). A psycholinguistic perspective on code-switching: Lexical, structural, and socio-interactive processes. In G. Stell & K. Yakpo (Eds.), *Code-switching between structural and sociolinguistic perspectives* (pp. 39–64). Berlin: De Gruyter.

Kroll, J. F., Bobb, S., & Wodniekca, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, *9*, 119–135.

Lee, C. C. P. (2005). The role of code-switching in a communications skills classroom. Unpublished M.A. (Applied Linguistics) thesis, Nanyang Technological University, Singapore.

Leimgruber, J. R. E. (2011). Singapore English. *Language and Linguistics Compass*, *5*, 47–62.

Lust, B., & Chien, Y.-C. (1984). The Structure of Coordination in First Language Acquisition of Chinese: Evidence for a universal. *Cognition*, *17*, 49–83.

Lust, B., Flynn, S., & Foley, C. (1996). What children know about what they say: Elicited imitation as a research method for assessing children's syntax. In D. McDaniel, C. McKee and H. S. Cairns, eds., *Methods for Assessing Children's Syntax*. Cambridge, MA: MIT Press.

Lust, B., Flynn, S., Blume, M., Park, S.W., Kang, C., & Kim, A. (2016). Assessing child bilingualism: Direct assessment of bilingual syntax amends caretaker report. *International Journal of Bilingualism*, *20*, 153–172.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language & Cognition*, *6*, 97–115.

Marian, V., Blumenfeld, H. K., Mizrahi, E., Kania, U., & Cordes, A. – K. (2013). Multilingual Stroop performance: Effects of trilingualism and proficiency on inhibitory control. *International Journal of Multilingualism*, *10*, 82–104.

MacLeod, C.M., & MacDonald, P.A. (2000). Interdimensional interference in the Stroop effect: uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences*, *4*, 383–391.

Meisel, J. (1994). Code-switching in young bilingual children: The acquisition of grammatical constraints. *Studies in Second Language Acquisition*, *16*, 413–419.

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*, 25–40.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*, 134–140.

Muysken, P. (2000). *Bilingual speech: A typology of codemixing*. Cambridge: Cambridge University Press.

Nicoladis, E., & Genesee, F. (1997). Language development in preschool bilingual children. *Journal of Speech-Language Pathology and Audiology*, *21*, 258–270.

Perez-Leroux, A. T., O'Rourke, E., & Gretchen, S. (2014). Language Dominance and Codeswitching Asymmetries. In MacSwan, J. (ed). *Grammatical Theory and Bilingual Codeswitching*. MIT Press. Pp 283–312.

Poarch, G. J., & Van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, *113*, 535–551.

Prior, A., & Gollan, T. H. (2011). Good language-switchers are good task-switchers: Evidence from Spanish–English and Mandarin–English bilinguals. *Journal of the International Neuropsychological Society*, *17*, 682–691.

Rosselli, M., Ardila, A., Santisi, M. N., Arecco, M. D. L., Salvatierra, J., Conde, A., & Lenis, B. (2002). Stroop effect in Spanish-English bilinguals. *Journal of the International Neuropsychological Society*, *8*, 819–827.

Rubdy, R. (2007). Singlish in school: Impediment or a resource? *Journal of Multilingual Matters*, *28*, 308–324.

Soveri, A., Rodriguez-Fornells, A., & Laine, M. (2011). Is there a relationship between language switching and executive functions in bilingualism? Introducing a within-group analysis approach. *Frontiers in Psychology*, *2*, 183–190.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Tan, Z. L. (2004). Xin jia po hua xiao xue sheng yu ma hun yong ge an yan jiu (A case study of code-mixing among Chinese primary school students in Singapore). Unpublished B.A. (Honors) thesis, Nanyang Technological University, Singapore.

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, *18*, 3–24.

Virtual Linguistics Lab Child Multilingualism Questionnaire(2016). Prepared by the Cornell Language Acquisition Lab and Virtual Center for Language Acquisition Members and Affiliates. In Blume, M., & B. Lust (eds) with Chien, Y., Dye, C., Foley, C., & Kedar, Y. *Research Methods in Language Acquisition: Principles, Procedures and Practices*. Mouton de Gruyter and American Psychological Association, Washington D.C.

Xu, D., Chew, C. H., & Chen, S. (1998). Language use and language attitudes in the Singapore Chinese community. In S. Gopinathan, A. Pakir, W. K. Ho, & V. Saravanan (Eds.), *Language, society and education in Singapore* (2nd ed., pp. 133–154). Singapore: Eastern Universities Press.

Yim, O., & Bialystok, E. (2012). Degree of conversational code-switching enhances verbal task switching in Cantonese-English bilinguals. *Bilingualism: Language and Cognition*, *15*, 873–883.

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press

Yow, W.Q., & Li, X. (2015). Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: Why variations in bilingual experiences matter. *Frontiers in Psychology*, *6*, 164.

Yow, W. Q., Patrycia, F., & Flynn, S. (2016). Code-switching in Childhood. In S. Montarnari & E. Nicoladis (eds). *Lifespan Perspectives on Bilingualism*.

Yow, W. Q., Tan, J. S. H., & Flynn, S. (2017). Code-switching as a marker of linguistic competence in bilingual children. *Bilingualism: Language and Cognition, First View article*.

Zied, K.M., Philippe, A., Pinon, K., Havet-Thomassin, V., Ghislaine, A., Roy, A., & Le Gall, D. (2004). Bilingualism and adult differences in inhibitory mechanisms: Evidence from a bilingual Stroop task. *Brain and Cognition*, *54*, 254–256.

https://doi.org/10.1017/S1366728918000299 Published online by Cambridge University Press