Bilingualism: Language and Cognition

cambridge.org/bil

Modulating bilingual language production and cognitive control: how bilingual language experience matters

Xuran Han[1,2,4], Li Wei[3,4] and Roberto Filippi[3,4]

[1]Institute for Lifecourse Development, School of Human Sciences, Faculty of Education, Health and Human Sciences, University of Greenwich, London, UK; [2]Centre for Research in Language and Heritage, Institute for Inclusive Communities and Environments, University of Greenwich, London, UK; [3]Institute of Education, UCL's Faculty of Education and Society, University College London, London, UK and [4]MULTAC (Multilanguage and Cognition Lab), Institute of Education, University College London, London, UK

Research Article

Cite this article: Han, X., Li, W., & Filippi, R. (2025). Modulating bilingual language production and cognitive control: how bilingual language experience matters. *Bilingualism: Language and Cognition* **28**, 120–134. https://doi.org/10.1017/S1366728924000191

Received: 30 May 2022
Revised: 31 January 2024
Accepted: 2 February 2024
First published online: 19 April 2024

Keywords:
bilingual language experience; code-switching production; cognitive control; L2 environment; language entropy

Author for correspondence:
Xuran Han;
Email: xuran.han.17@ucl.ac.uk

Abstract

The Adaptive Control Hypothesis and the Control Process Model propose that bilingual language use in different interactional contexts requires control processes that can adapt in different ways to linguistic demands. This study explored the effects of language experience on cognitive flexibility and inhibition among 41 Chinese–English bilingual adults. In particular, it aimed to investigate the relationship between spontaneous language production (i.e., bilingual conversation and narration tasks) and cognitive control. Participants' inhibitory control and cognitive flexibility efficiency was measured through verbal and spatial Stroop tasks, and a colour-shape switching task. Overall, it showed that frequent practices of intersentential switching in speech production resulted in significant facilitatory effects in both verbal and nonverbal inhibitory control. This study provides new evidence for the importance of bilingual language experience in adaptive cognitive control in naturalistic speech production and furthers our theoretical knowledge of the relationship between the language system and crucial domain-general cognitive processes.

Introduction

Recent research highlighted the significance of individual difference in bilingual language experience and its effects on cognitive control development (e.g., DeLuca et al., 2019; Ooi et al., 2018; Surrain & Luk, 2019). For example, numerous studies have started to take into account language experience-related factors, such as L2 proficiency, habitual language use context and L2 immersion, and to investigate the interconnections between bilingual language experience and cognitive control (e.g., Beatty-Martínez et al., 2020; Han et al., 2022; Singh & Kar, 2018). Furthermore, bilingualism is increasingly discussed as a dynamic and interactive life experience with internal and contextual features (Grosjean, 2013; Surrain & Luk, 2019). A continuum of complex aspects in the individual bilingual experience trajectory may confer different outcomes on domain-general cognitive control development. Therefore, the relationship between bilingualism and cognitive control cannot be oversimplified by binary comparisons between bilinguals and monolinguals.

In addition to empirical studies, relevant theoretical frameworks, the Adaptive Control Hypothesis (Green & Abutalebi, 2013) and the Control Process Model (Green & Li, 2014), postulated the interactive relationship between the deployment of cognitive control and bilingual language use experience. They suggested that bilinguals' cognitive control dynamically adapts to differences in interactional contexts that they habitually engage in and patterns of code-switching that they habitually produce in natural communication. However, fewer studies have explored cognitive control in bilinguals' code-switching in naturalistic settings or spontaneous language production by addressing its interaction with their language use habits. This study, therefore, aims to adopt a more comprehensive approach to bilingual participants' habitual language use practices, and explore how individual difference in bilinguals' language use habits affects their spontaneous bilingual language production and efficiency in cognitive control.

Code-switching production and cognitive control

Code-switching is commonly used in bilingual language practices and imposes high cognitive demands due to coordinating two languages in one integral utterance. Utterances with mixing

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

of languages can be classified as intersentential switching and intrasentential switching based on Poplack's (1980) typology. Intersentential switching, as measured in the present study, referred to bilinguals switching between two languages across clauses or sentences (e.g., "我一直努力地工作挣钱 [I always work hard to make more money]; and this is the only thing I can do."). In this example, the first clause is Chinese, and the second clause is in English, representing an intersentential code-switching pattern. According to Muysken (2000), intrasentential switching manifests in three different ways: insertion, alternation and congruent lexicalisation. In insertional switching, words or lexical constituents from both languages are co-used and switched within one single utterance (e.g., "我必须在 online meeting 前把report 发给他。 [I have to email him the report before the online meeting]"). Chinese and English co-exist in this sentence; specifically, English word chunks are embedded in the Chinese sentential framework. Alternational switching happens directionally in the middle of a sentence, i.e. the sentence begins in a stretch of words in one language and ends in a different one. For example, 雪下的太大了， 我的 train got severely delayed [It snowed so heavily that my train got severely delayed].

In utterances with congruent lexicalisation, different languages are actively intertwined at both the grammatical and words/morphemes levels to achieve a shared language structure. However, congruent lexicalisation is not a commonly produced code-switching pattern between two typologically distant languages such as Chinese and English. Furthermore, in this study, insertion was found to be the predominant code-switching pattern in the participants' language production. Therefore, this study analysed participants' insertion practices, as one representative type of intrasentential switching[1], and participants' intersentential switching practices in the language production tasks.

Numerous studies have explored the modulation effects of code-switching practices on bilinguals' cognitive control performance; however, inconsistent results were reported by studies focusing on non-cued code-switching production (e.g., Gollan et al., 2014; Kleinman & Gollan, 2016) and those focusing on cued code-switching production (e.g., Acheson et al., 2012; Meuter & Allport, 1999). The picture-naming task is widely adopted in these studies to measure participants' code-switching. Participants in the picture-naming task are instructed to name the presented objects in different languages. While this task assessed bilinguals' code-switching at the level of nouns and measured participants' performance difference between switch and non-switch naming trials, it did not sufficiently capture the multifaceted nature of code-switching behaviours that bilinguals exhibit in their daily communication. In naturalistic settings, bilinguals can produce longer stretches of utterances that incorporate different patterns of code-switching, such as language switching at clause and verbs levels. Therefore, to gain a more comprehensive understanding of how bilinguals habitually combine their two languages and engage in code-switching during spontaneous language production, it is essential to employ more elaborate measures to assess and quantify the various dimensions of code-switching in bilingual communication.

Pausing in bilinguals' spontaneous language production was measured as an important index related to oral fluency of their bilingual speech (Green & Li, 2014). Pauses in speech have been discussed as hesitation and monitoring phenomena (Hlavac, 2011), conveying communicative signals associated with not only speech planning (Green & Li, 2014) but also closely linked with cognitive processing (Gardner-Chloros et al., 2013; Schilperoord, 2002). Navracsics (2004) calculated the number of code-switches, errors, silent and filled pauses in participants' bilingual and monolingual speeches. The results showed that silent and filled pauses are the most common dysfluencies in participants' bilingual and monolingual speeches, with a higher frequency observed in their bilingual speeches. Gardner-Chloros et al. (2013) conducted a detailed analysis on pauses in participants' naturalistic speech to explore the connections between pausing and code-switching in London and Cyprus bilingual groups. However, they did not find an association between code-switching and pausing frequency in speech.

Noticeably, bilinguals' language proficiency can also affect the frequency of pausing in spontaneous speech and their utterance fluency (e.g., Révész et al., 2016); therefore, the association between participants' L2 proficiency and spontaneous language production fluency is also analysed in this study.

Bilingual language use experience in cognitive control modulation

In the Adaptive Control Hypothesis (ACH), Green and Abutalebi (2013) discussed how bilinguals' language use and cognitive control demands varied across three interactional contexts: single-language, dual-language and dense code-switching contexts. In the single-language context, languages are used separately in distinct situations (e.g., use Chinese at home but use English at work). Code-switching seldom happens in this context; therefore, bilinguals must constantly monitor linguistic competition and efficiently control lexical interference from their co-activated languages to ensure successful single-language production in a given situation. Cognitive processes, including goal maintenance and interference control, are supposed to be actively involved in motivating bilinguals' language production in this context. In dual-language contexts, bilinguals use their languages alternately in one situation or with different interlocutors during communication. For example, Chinese–English bilinguals communicate in Chinese with Chinese interlocutors while they will switch to speak in English when English interlocutors join the conversation. Bilinguals are expected to have intensive intersentential switching practices in the dual-language contexts. Bilinguals in dense code-switching contexts alternate frequently and rapidly between two languages, often within a single sentence or utterance. They seamlessly incorporate words or phrases from both languages, creating a unique linguistic blend or hybrid. Therefore, language task schemas co-operate with each other, and the cognitive processes of interference control and goal maintenance are supposed to be less engaged when bilinguals routinely mix two languages within utterances in the dense code-switching contexts (Green & Abutalebi, 2013).

In the Control Process Model (CPM), Green and Li (2014) emphasised the variations of cognitive control demands in processing different code-switching patterns. That is, the degree of activation of lexical items and the grammatical frameworks across two languages leads to the deployment of different cognitive control strategies in producing various code-switching patterns (Ng & Yang, 2021; Treffers-Daller, 2009). For example, bilinguals use the language that comes most readily to them to formulate their intrasentential switching utterance in dense code-switching contexts. Increasing demands of opportunistic planning in language control processes are supposed to be involved in such contexts to facilitate bilingual speakers' dense code-switching production. Therefore, frequently engaging in the dense code-switching communication may largely exercise bilinguals' opportunistic planning processes, but with less exercise on inhibitory control. In contrast, bilinguals need to constantly monitor and control co-activated languages to maintain linguistic constituents from the two languages as mutually independent without structural adaption during intersentential switching production (Green & Li, 2014). As a result, intersentential switching is more cognitively demanding and bilinguals' efficiency in more aspects of cognitive control processes, such as inhibitory control, conflict monitoring and cue detection, is expected to improve through frequent intersentential switching practices.

Although numerous studies examining the ACH and CPM have reported positive evidence for the varied degree of modulation effects of different interactional contexts or code-switching patterns on bilinguals' cognitive control (e.g., Choo et al., 2021; Y. J. Wu & Thierry, 2013), some studies (e.g., Hartanto & Yang, 2020; Kałamała et al., 2020) reported inconsistent results. For example, Blanco-Elorrieta and Pylkkänen (2017) conducted a MEG study to compare bilingual code-switching performance in multiple interactional contexts, ranging from an artificial switching condition to a fully naturalistic conversation condition. In their study, 19 Arabic–English bilinguals were instructed to name the picture in a specific language according to the rules in a bilingual-interlocutor-context, a monolingual-interlocutor-context and a laboratory-colour-cued condition. They found that bilinguals' voluntary switching in a fully natural conversation scenario did not engage significant switch costs, as compared to switching languages in the laboratory-colour-cued condition. Their results reflected that processing code-switching in naturalistic interactional contexts was relatively "effortless" or less cognitively demanding for bilinguals. Besides, their study highlighted the necessity of discussing the cognitive control involved in bilingual language production in naturalistic conditions. However, it is still challenging to measure and quantify bilinguals' habitual code-switching practices and interactional context involvement in daily life. Furthermore, L2 environment immersion (e.g., Bonfieni et al., 2019; Heidlmayr et al., 2014), L2 proficiency and code-switching frequency (e.g., Beatty-Martínez & Dussias, 2017; Kheder & Kaan, 2021) within bilingual language experience are interconnected and exert long-term effects on the development of bilinguals' cognitive control abilities.

The present study

The study aimed to understand the interactive effects of multifaceted factors in bilingualism on bilinguals' language production and cognitive functions. It investigated how bilingual language use habits affected bilinguals' spontaneous language production and domain-general cognitive control performance among a group of Chinese–English bilinguals living in the L2 environment. Specifically, the current study addressed:

a) What is the relationship between bilinguals' spontaneous language production performance and their performance in verbal and nonverbal inhibitory control tasks?
b) How does bilinguals' spontaneous language production performance associate with their nonverbal shifting performance?
c) How does bilinguals' habitual language use experience, especially L2 proficiency, code-switching frequency and L2 environment immersion, affect their spontaneous language production and cognitive control performance?

It predicted that:

i) Participants who habitually produce intersentential switching in the spontaneous language production will outperform those who seldomly switch, in the inhibitory control tasks.
ii) Bilinguals' intersentential switching frequency in language production tasks will positively associate with their shifting efficiency.
iii) Participants with higher L2 proficiency will show higher shifting efficiency.
iv) Habitual single-language context bilinguals will show reduced efficiency in both verbal and nonverbal inhibitory control tasks. Additionally, increased immersion in the L2 environment will positively affect bilinguals' inhibitory control and shifting efficiency.

Method

The study obtained approval from the Ethics Committee of the Institute of Education, University College London, for empirical research involving human participants. Before data collection, participants received a brief explanation of the study's instructions. Each interested participant received an information sheet and a consent form, and only those who signed the consent form were involved in the study.

Participants

Forty-one (9 males; Mean age = 26, SD = 2.92, range: 21-33 years old) right-handed healthy Chinese–English bilingual adults residing in English-speaking countries, such as the UK, America, Australia, Canada, and Ireland, participated in the study. They were raised in Chinese Mandarin-speaking families, received education in China, and later relocated to English-speaking countries for higher education or work in adulthood. They were all late bilinguals, who learnt English as their L2 after their native language, Chinese Mandarin, was well-acquired (Mean L2 AoA = 6.7 years old). At the time of this study, participants had lived in English dominant communities (e.g., the UK, America) for 3.5 years on average.

The Lexical Test for Advanced Learners of English (LexTALE test, Lemhöfer & Broersma, 2012) was used to measure participants' L2 proficiency. This standardised test is in an un-speeded visual lexical decision task for assessing bilinguals' L2 proficiency. The higher score participants got indicates their higher L2 proficiency level. Overall, they obtained moderately-high proficiency in English (Mean = 63.75%).

Participants' self-reported language proficiency levels, habitual bilingual language use patterns, and code-switching habits were measured through LEAP-Q (Marian et al., 2007), a code-switching and interactional context questionnaire (Hartanto & Yang, 2016), and a Chinese-translated bilingual switching questionnaire (BSWQ) (Rodriguez-Fornells et al., 2012), respectively.

Participants' habitual bilingual language use was quantified through computing language entropy (Gullifer & Titone, 2020) in four different communicative contexts: home, workplace, school and social activities. In this study, participants' language entropy value in each context ranges from 0 to 1, with a higher entropy value indicating a more balanced and diverse use of two languages in that context (Gullifer & Titone, 2020). A language entropy value of 0 indicates that only one language is used in that context. Conversely, the maximum language entropy value is reached when two languages are used and combined in a highly balanced and diverse manner in that context. Table 1 shows information related to bilingual participants' language use experience.

https://doi.org/10.1017/S1366728924000191 Published online by Cambridge University Press

Table 1. Demographic, language history and bilingual language use information of the Chinese-English bilingual participants

	Sample (N = 41)	
	Mean	SD
Basic demographics		
Age (years)	26.17	2.92
L2 AoA (years)	6.68	3.66
LexTALE score	63.75	9.52
L2 exposure (years)	3.49	2.29
Chinese use in bilingual conversations (%)	67.32	20.86
English use in bilingual conversations (%)	32.68	20.86
Proportion of reading in Chinese (%)	61.34	23.37
Proportion of reading in English (%)	38.66	23.37
Accent perception in L2		
Self-perceived (1-10)	4.17	2.18
Other-recognised (1-10)	5.22	3.37
Self-reported L2 proficiency		
Reading (1-7)	5.71	0.98
Comprehension (1-7)	5.54	0.93
Writing (1-7)	4.93	1.01
Speaking (1-7)	5.22	1.01
Self-reported L1 proficiency		
Reading (1-7)	6.80	0.40
Comprehension (1-7)	6.98	0.15
Writing (1-7)	6.34	0.72
Speaking (1-7)	6.85	0.15
Bilingual switching habits		
L1 switch tendencies	7.98	2.22
L2 switching tendencies	8.32	1.96
Contextual switch	8.71	2.73
Unintended switch	7.88	2.35
Habitual code-switching and interactional context		
Single-language score	69.49	20.91
Dual-language score	4.83	1.86
Intersentential switching index	2.16	0.75
Intrasentential switching index	2.78	0.95
Language entropy		
Home	0.49	0.33
School	0.54	0.34
Work	0.31	0.40
Social activities	0.91	0.31

Materials and task design

Due to the pandemic lockdown, this study was conducted remotely. Behavioural tasks were generated using PsychoPy (Peirce et al., 2019) and conducted via Pavlovia (http://pavlovia.org/). Language production tasks were conducted through online one-to-one meetings between the researcher and participants.

Spontaneous language production tasks

A naturalistic conversation task and a story narration task were used to capture participants' spontaneous bilingual language production patterns. In the conversation task, participants could switch between Chinese and English freely. In the story narration task, participants should recount stories based on provided pictures using either one language or both languages. Both tasks aimed to analyse pauses, monolingual utterances and frequencies of different code-switching utterances produced by the participants.

1. Naturalistic conversation task. This task is designed as a semi-structured conversation in which participants discussed their weekend plans and favourite weekend activities with the Chinese–English bilingual researcher. Initially, participants were given five minutes to tell the researcher what they did for last weekend. Participants in the monologue were free to switch between Chinese and English in their habitual way. The instructions for this task were communicated in Mandarin. Accordingly, the researcher asked three interactive questions to participants. One question was asked in English, one in Chinese and one with Chinese–English code-switching. This is to prompt participant's language production and assess whether they used two languages differently when responding to the questions asked in a monolingual or bilingual manner. Participants were able to use the two languages in their preferred way to answer the questions just as they were in a conversation with their bilingual friends. The order of the three questions was counterbalanced across participants and presented after the five-minute monologue (see Supplementary Table 1 for details).

2. Story narration task. Adapted from the story recounting task and picture descriptive task used by Toribio (2001) and Lloyd-Smith et al. (2020), this self-designed bilingual speech production task aimed to engage participants in spontaneous language production and measure their code-switching in their narrative speech for different stories. Different from the conversation task, this narration task controlled the dysfluency in participants' speech caused by memory demands on recalling and redescribing their past experience. Moreover, this task constrained the variation in topics participants might produce in their speech, thus reducing confounds in narrative data analysis caused by different familiarity and variation in speech topics (Zantout, 2019).

Participants in this task were given a set of pictures without any linguistic cues, and were instructed to recount a story illustrated by those pictures in five minutes. This task was administered in single-language contexts in which participants were instructed to use one specified language, and a bilingual context in which they were able to code-switch in a voluntary way. Specifically, they were instructed to produce two monolingual narrations (i.e., Chinese-only and English-only) for each set of pictures designed for single-language production (i.e., Hua Mulan and the little match girl); and they should narrate a bilingual story with code-switching based on a set of pictures of three little pigs.

Given that bilinguals' naturalistic language interaction was influenced by the interlocutors and the interactional context, the researcher in the monolingual narration condition used Chinese and English accordingly to specify the language for

https://doi.org/10.1017/S1366728924000191 Published online by Cambridge University Press

participants to use in narration. Bilingual instruction was used in the bilingual language condition to require participants to tell a story with two languages combined.

The order of the single-language story narration was counterbalanced across participants, but bilingual story narration was always conducted last. Short breaks were offered after each narration completion, and participants' narrative speech in this task was recorded.

Verbal Stroop task

This task was adapted from the computerised Stroop (1935) colour-naming task to measure participants' verbal inhibitory control performance. Chinese words were used as stimuli as the participants' native language is Chinese. There were four types of trials based on four different colours: red, blue, yellow and green, in this task.

In this task, participants saw four colour words in Chinese (i.e., 红(red), 绿(green), 蓝(blue) and黄(yellow)) printed in four ink colours. Their task was to respond to the ink colours of the Chinese words by pressing the corresponding keys. In congruent trials, the meaning of the colour word matches with the ink colour in which the word was displayed (e.g., '红' displayed in red). For incongruent trials, the colour word's meaning mismatched with the ink colour of the word (e.g., '黄' displayed in green). There were 24 neutral trials, which included four Chinese non-colour words: 吃(eat), 忙(busy), 路(road), 富(rich), printed in four different ink colours. A practice session with 12 trials was presented at the beginning; afterwards, 120 trials were included in five formal experimental blocks. Each trial began with a centred fixation cross (+) for 500ms and followed by a stimulus that remained on the screen for 3,500ms or until a response happened. A blank black sheet presented for 300ms immediately after a response or stimulus disappeared.

Participants firstly completed the neutral block. Then, they completed a congruent and an incongruent block, respectively, and each of them including 24 trials. The order of these two blocks was counterbalanced across participants. Participants completed the mixed block of 48 trials with an equal number of congruent and incongruent trials at last. Trials in the mixed block were presented in a fixed pseudo-random order. Participants' reaction time (RT) and response accuracy in this task were automatically recorded for data analysis.

Spatial Stroop task

The spatial Stroop task, adapted from Blumenfeld and Marian (2011, 2013), was to measure participants' nonverbal inhibitory control performance. Participants in this task should judge the arrow's direction while ignoring the location of the arrow as quickly and accurately as possible. The trial was congruent when the arrow's pointing direction matched with its location on the screen (e.g., left-pointing arrow presented on the left side of the screen); otherwise, when the arrow's pointing direction was inconsistent with its location, it was an incongruent trial (e.g., up-pointing arrow presented at the bottom of the screen). Each trial began with a centred fixation cross for 500ms, followed by a 3,500ms presentation of the stimulus. After that, a blank sheet immediately appeared for 300ms.

The task included one neutral block, in which a circle was presented as a visual stimulus on the screen. A mixed block, containing equal numbers of congruent and incongruent trials, was presented at the end of this task. All visual stimuli can appear in any of four locations (top, bottom, left and right) on the screen.

After completing 12 practice trials at the beginning, participants should complete 24 neutral trials. Congruent and incongruent block with 24 trials respectively followed, and the order of the two blocks was counterbalanced across participants. The mixed block, containing 48 trials with equal numbers of congruent and incongruent trials, was completed at last. Trials in the mixed block were presented in a fixed pseudo-random order.

Colour-Shape switching task

The colour-shape switching task was adapted from Yim and Bialystok (2012) and Barac and Bialystok (2012) to measure participants' domain-general shifting ability. Participants in this task saw three pictures at a time presented on the screen. Two target pictures (an orange rabbit and a green duck) were presented at the top of the screen while one stimulus picture, either an orange duck or a green rabbit, appeared below the two targets.

Participants should match the stimulus picture to one of the target pictures, according to a cue that appeared with the stimulus picture. A jigsaw was the cue for matching based on same shape, while a palette indicated matching the stimulus with a target based on same colour. If the stimulus matched with the left side target picture, participants should press the left arrow on the keyboard to respond; otherwise, they should press the right arrow.

Each trial started with a 250ms fixation cross at the centre of screen, followed by picture stimuli that remained on the screen for 3,000ms or until a response was made. Following the response, the next trial started after a delay of 1,000ms.

The task comprised 200 trials distributed across two single-task blocks (25 colour trials and 25 shape trials) and three mixed blocks. In single-task blocks, participants consistently matched stimulus and target pictures based on either the same shape or colour. In mixed blocks, successive trials could have the same matching criteria (repeated trials) or different criteria (switch trials), with a 1:1 ratio. Following an 8-trial practice block, formal blocks were administered in a fixed order, beginning with two single-task blocks and followed by three mixed blocks. The order of the single-task and mixed blocks was counterbalanced. Participants were instructed to perform the task quickly and accurately, with both RTs and response accuracy recorded. The general procedure and design of the three cognitive control tasks are illustrated in Figure 1.

Procedure

All participants were invited to join the study remotely in their quiet rooms. Initially, an email with a meeting schedule invitation was sent for participants to confirm their availability. Those who promptly confirmed were invited to a pre-arranged online meeting where two spontaneous language production tasks were conducted. Following the meeting, participants received links to three online cognitive tasks, the LexTALE test, and the questionnaires. They had the flexibility to complete these tasks in any order, but the cognitive tasks were consistently completed after the language production tasks. The study procedure is illustrated in the Figure A.1. in Appendix 1.

Statistics

Data collection

For each participant, the naturalistic conversation task yielded output in four conditions: a 5-minute monologue, responses to

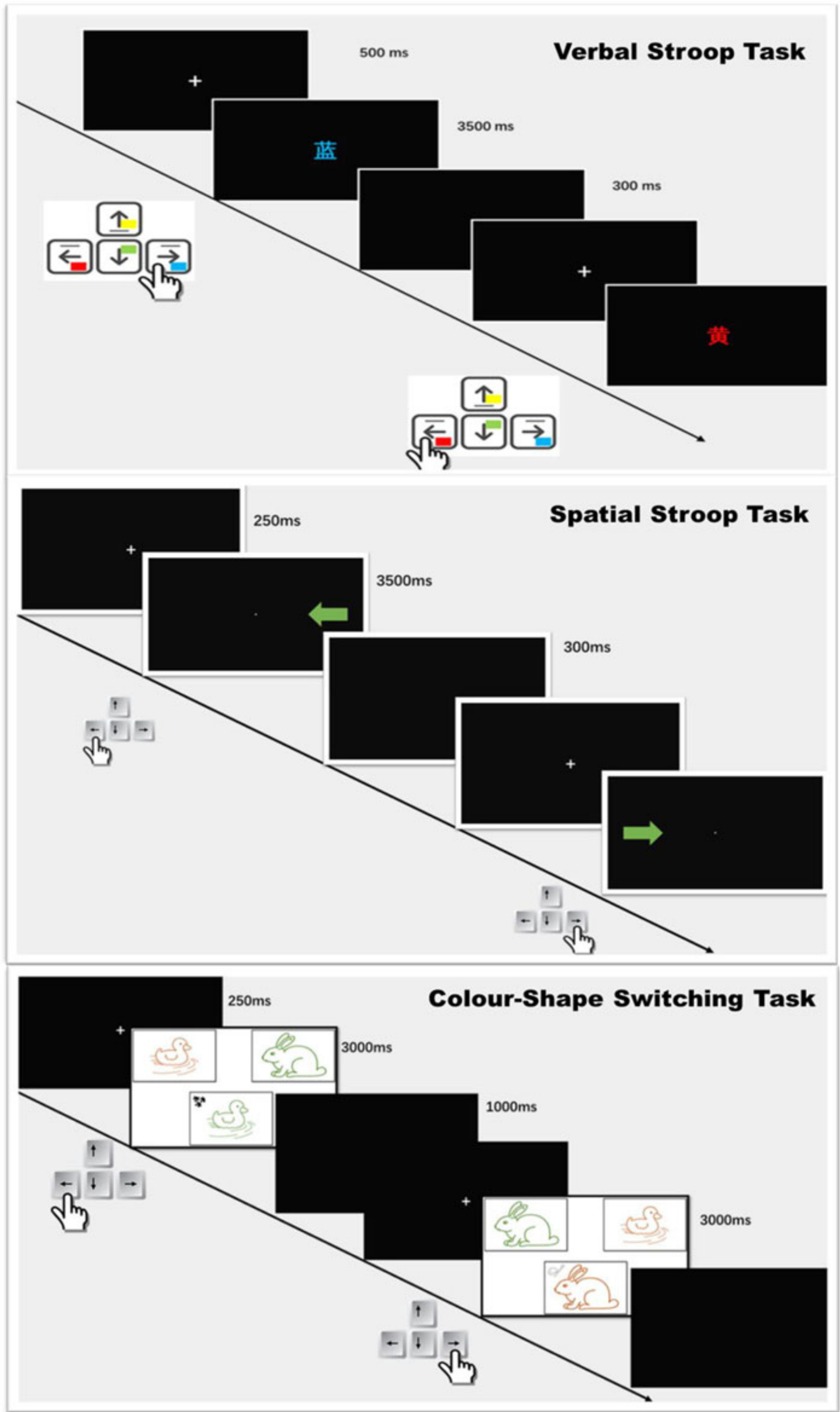

Figure 1. Illustrations of the trial design and procedure for the three cognitive control tasks used in this study.

Mandarin and English questions, and responses to a bilingual question. In the story narration task, participants provided speech outputs in English-only, Chinese-only, and bilingual narratives. Audio recordings and transcripts of language production tasks were analysed using Praat phonetic software (version 6.1.24, Boersma & Weenink, 2020). The software's internal *textgrid* script sliced each recording into "sounding" segments (speech) and "silent" segments (pauses). Notably, pauses included both silent and filled pauses with meaningless lexicalizations like 'uh', 'eh' or 'um'. To ensure accurate counting, all sounding segments, including pauses, were manually checked, as filled pauses were sometimes misidentified as sounding segments by Praat.

Any pauses over 250ms produced by participants in the two tasks were identified for pause frequency calculations. Participants producing longer speech are supposed to have a higher number of pauses; therefore, ratios of their pause frequency (ϕ) in these two tasks were computed, following Zantout's (2019, p. 272) study: $\phi = \frac{n}{t_{speech} - t_{pause}}$, where n is the total number of pauses over 250 ms, t_{speech} the total speech length (in seconds) and t_{pause} the total duration of pauses (in seconds) in the speech. Each participant's mean pause duration ($t_{pause,avg}$) was computed as $t_{pause,avg} = \frac{t_{pause}}{n}$.

The total number of utterances for each participant's speech in the two tasks was counted. Additionally, the frequencies of intersentential and intrasentential switching utterances were calculated respectively, in the conversation task. However, participants were only able to freely switch between languages while narrating the story of the three little pigs; therefore, code-switching frequency in their speech was only calculated in this story narration. The number of utterances which contained only English or Chinese was calculated as the frequency of English/ Chinese in participants' speech samples and in the story narration task. To control for bias in code-switching frequency calculations stemming from variations in utterance numbers across participants, percentages of code-switching and monolingual utterances were computed for each participant in each speech. For instance, dividing the number of intersentential switching utterances by the total number of utterances in the conversation task yields the participant's percentage of intersentential switching. The same computation method applies to percentages of monolingual utterances in speech samples. Supplementary Table 2 showed the variables measured in the two language production tasks and indexes computed for subsequent analyses.

Participants' RTs (ms) and their response accuracy in the three cognitive tasks were collected and calculated for subsequent analysis.

Data preparation and analysis

Initial pauses in language production tasks were eliminated, as they were typically attributed to participants' preparatory actions, such as clearing their minds before speaking, rather than actual language production challenges. Furthermore, pauses arising from microphone adjustments during speeches are excluded, as they were not related to participants' language production difficulties.

In the verbal Stroop task, data from 4,920 trials were collected, with 17 missing values (0.35%) excluded from analysis. Trials with RTs below 200ms or above 3,500ms were removed, along with values beyond 2.5 standard deviations from participants' individual mean RTs (n = 153). After this data preprocessing, 4,750 data points remained. Analysis focused on correct response trials, excluding 1,002 trials with incorrect responses, resulting in 3,816 trials for participants' RTs analysis.

Participants' responses to 4,920 trials in the Spatial Stroop task were recorded, with no missing data observed. Three values below 200ms were eliminated, and 13 values above 2.5 standard deviations of participants' individual mean RTs were excluded from analysis. After removing 127 trials with incorrect responses, 4,676 data points remained for participants' RT analysis.

A total of 7,872 responses were collected in the colour-shape switching task, with. 99 missing values (1.26%) removed from the dataset. The following responses were also excluded: correct responses with RTs below 200ms (n = 2), and values above 2.5 standard deviations of participants' individual mean RTs (n = 211). Incorrect response trials (n = 493) were also excluded from participants' RTs analyses, having 7,067 trials remained.

Linear mixed effects models in R (Version 4.0.2; R Studio Team, 2020) were used to analyse participants' task performance in the three cognitive tasks. For the analysis of RTs, a mixed model was run, using the lmer function as implemented in the lme4 package for R (Version, 1.1 - 26; Bates et al., 2015). Participants' response accuracy in these tasks was analysed through the generalised linear mixed effects model with a logistic link function. The model was run with a glmer function as implemented in the lme4 package for R (1.1.21; Bates et al., 2015). The random effect of subjects was included in the analyses to account for variability across participants. Reported p-values were calculated based on Satterthwaite's method as implemented in the lmerTest package in R (Kuznetsova et al., 2017). Interactions between participants' language production and their domain-general cognitive control performance were examined through linear mixed effect models to associate participants' cognitive task performance and their z-scored bilingual language use habits, language entropy in four different contexts and their language production task performance.

Results

Spontaneous language production task performance

Participants' mean pause duration, pause frequency ratio and percentages of code-switching in the naturalistic conversation task and the story narration task were analysed. Descriptive statistics of participants' performance in the two tasks are presented in Supplementary Table 3.

Pause frequency

ANOVAs revealed that participants' pause frequency in bilingual speech varied significantly across the conversation and narration tasks, F (5, 240) = 15.99, $p < .001$, $\eta^2 = .25$. Specifically, participants' pause frequency ratio in bilingual narration was significantly higher than in conversational speech ($t = 7.16$, Cohen's $d = 1.60$, $p < .001$). Moreover, participants paused more and showed less fluency in bilingual narration than in follow-up question answering in the conversation task ($p < .001$). However, participants' pausing frequency in conversational monologue and follow-up question answering was comparable, F (4, 200) = 15.99, $p = .81$, $\eta^2 = .01$.

Participants' pause frequencies between monolingual and bilingual story narrations were significantly different, F (4, 200) = 14.38, $p < .001$, $\eta^2 = .22$. Participants' pause frequency was positively correlated with their language proficiency levels, i.e., more pauses during English narration, regardless of the story's cultural background. Specifically, as compared to their English narrations, participants performed more fluently in narrating the little match girl story ($t = -5.11$, Cohen's $d = -1.18$, $p < .001$) and Hua Mulan story ($t = -4.97$, Cohen's $d = -1.13$, $p < .001$) in Chinese.

Moreover, participants paused less frequently in their Chinese narration of the little match girl story as compared to narrating the Hua Mulan story in English ($t = -5.37$, Cohen's $d = -1.28$, $p < .001$). A similar pattern was also found in the comparison of their Chinese narration of Hua Mulan and English narration of the little match girl story ($t = -4.71$, Cohen's $d = -1.04$, $p < .001$). Interestingly, participants' pause frequency in bilingual narration was lower than English monolingual narrations ($p < .001$); while being comparable to Chinese monolingual narrations ($p = 1.00$).

Pause duration

Results showed that mean pause duration in participants' speech varied across the two tasks, F (5, 240) = 7.45, $p < .001$, $\eta^2 = .13$.

Specifically, mean pause duration in participants' bilingual narration was longer than in their Chinese speech ($t = 4.97$, Cohen's $d = 1.51$, $p < .001$), English ($t = 3.23$, Cohen's $d = 0.71$, $p = .02$) as well as bilingual question answering ($t = 3.93$, Cohen's $d = 1.07$, $p = .002$). Participants' mean pause duration did not differ within their speech for different question answering. Furthermore, they had comparable mean pause duration ($t = 1.44$, Cohen's $d = 0.48$, $p = 1.00$) in bilingual narrative and conversational speech. However, their mean pause duration in different story narrative speech was comparable, F (4, 200) = 0.69, $p = .60$, $\eta^2 = .01$.

Code-switching frequency

In general, there was a significant difference in participants' percentages for code-switching frequency across the two tasks, F (19, 800) = 33.58, $p < .001$, $\eta^2 = .44$. Participants produced significantly more Chinese monolingual utterances as compared to any other types of utterances (i.e., English monolingual, intersentential switching and intrasentential switching) in the two tasks ($p < .001$). Fewer English monolingual utterances were found in participants' speech for post-conversation question answering as compared to their conversational speech ($p < .001$) and bilingual story narration ($p < .001$). Participants also produced intrasentential switching more frequently than intersentential switching in their conversational speech ($t = -5.12$, Cohen's $d = -2.45$, $p < .001$); and in bilingual narration ($t = -4.80$, Cohen's $d = -2.12$, $p < .001$).

Spearman's rank correlation analysis revealed a negative correlation between participants' intrasentential switching percentages in bilingual narrations and their index of intrasentential switching assessed through the code-switching and interactional context questionnaire (Hartanto & Yang, 2016), r_s (39) = -.36, $p = .02$. However, participants' intersentential switching percentages in bilingual narrations did not show significant correlation with their index of intersentential switching. The results showed that participants' self-reported code-switching frequency might not sufficiently reflect their code-switching frequency in spontaneous language production. The null correlation between intersentential switching frequency in bilingual narrations and self-reported questionnaire was reasonable, since these participants seldomly use intersentential switching in the narrative task.

Cognitive control tasks performance

RT(s) and response accuracy in the verbal Stroop task

Participants' RTs and response accuracy in the verbal Stroop task were analysed through linear mixed effects models. Table 2 presents their performance across three different blocks in this task.

Results showed that participants' RTs varied across different blocks, F (1, 39.82) = 11.55, $p = .002$, and different congruency, F (1, 39.45) = 111.76, $p < .001$. An interactive effect of block*congruency was found in participants' RTs, F (1, 37.13) = 6.65, $p = .01$. Figure 2 illustrates participants' RTs for congruent and incongruent trials across different blocks.

As expected, participants' incongruent RTs were greater than congruent RTs in the mix block (t (39.7) = -6.67, $p < .0001$) and single blocks, t (39.6) = -9.45, $p < .0001$. Participants' congruent RTs in the mixed block were greater than congruent RTs in the single block (t (39.6) = 4.92, $p = .0001$), but smaller than incongruent RTs in the single block (t (39.5) = -6.07, $p < .0001$). However, incongruent RTs in both the single and mixed blocks were comparable, t (39.6) = 0.71, $p = 1.00$.

Similarly, participants' response accuracy was affected by congruency ($\beta = -0.54$, SE = 0.08, $z = -6.65$, $p < .0001$), block ($\beta = -0.58$, SE = 0.08, $z = -7.08$ $p < .0001$) and their interactive effects ($\beta = 0.98$, SE = 0.16, $z = 5.98$, $p < .0001$). Figure 3 presents the verbal Stroop task response accuracy across different trials and blocks.

Participants performed more accurately for congruent trials than incongruent trials in the mixed block ($\beta = 1.03$, SE = 0.13, $z = 8.19$, $p < .0001$), and even more so than the congruent ($\beta = 1.07$, SE = 0.13, $z = 8.54$, $p < .0001$) and incongruent ($\beta = 1.12$, SE = 0.13, $z = 8.92$, $p < .0001$) trials in the single block. However, their response accuracy for congruent and incongruent trials in single blocks was comparable ($\beta = 0.06$, SE = 0.10, $z = 0.53$, $p = 1.00$). There was no significant difference in response accuracy between incongruent trials in the mix and single blocks ($\beta = 0.09$, SE = 0.11, $z = 0.86$, $p = 1.00$).

RT(s) and response accuracy in the Spatial Stroop task

Participants' RTs and response accuracy in the spatial Stroop task were analysed through linear mixed effects models. Table 3 shows their performance across different blocks in this task.

The results showed that participants' RTs varied across blocks (F (1, 40.06) = 4.23, $p = .046$) and congruency (F (1, 36.35) = 159.02, $p < .001$.). The interactive effect of block and congruency was found to affect participants' RTs in this task, F (1, 38.57) = 47.83, $p < .001$. Figure 4 showed RTs in incongruent and congruent trials across different blocks in the spatial Stroop task.

Results showed that congruent RTs in both single (t (39.7) = -11.09, $p < .0001$) and mix (t (39.3) = -5.71, $p < .0001$) blocks were significantly smaller than incongruent RTs in the mixed block. Moreover, in single blocks, participants' congruent RTs were significantly smaller than their incongruent RTs (t (39.9) = -11.86, $p < .0001$). In addition, participants' incongruent RTs in the mixed block were significantly smaller than in the single block (t (39.4) = -3.36, $p = .01$). However, congruent RTs in the mixed block were greater than in the single block (t (40) = 6.42,

Table 2. Participants' RTs and response accuracy in the verbal Stroop task

		RTs (ms)		Accuracy (%)	
Block	Trial	Mean	SD	Mean	SD
Neutral (Ne)	neutral	823.53	135.24	94.97	11.59
Single	congruent	712.83	122.91	72.52	8.72
	incongruent	895.45	173.99	71.12	9.26
Mix	congruent	787.72	146.22	88.39	9.55
	incongruent	899.51	159.04	73.20	9.52

https://doi.org/10.1017/S1366728924000191 Published online by Cambridge University Press

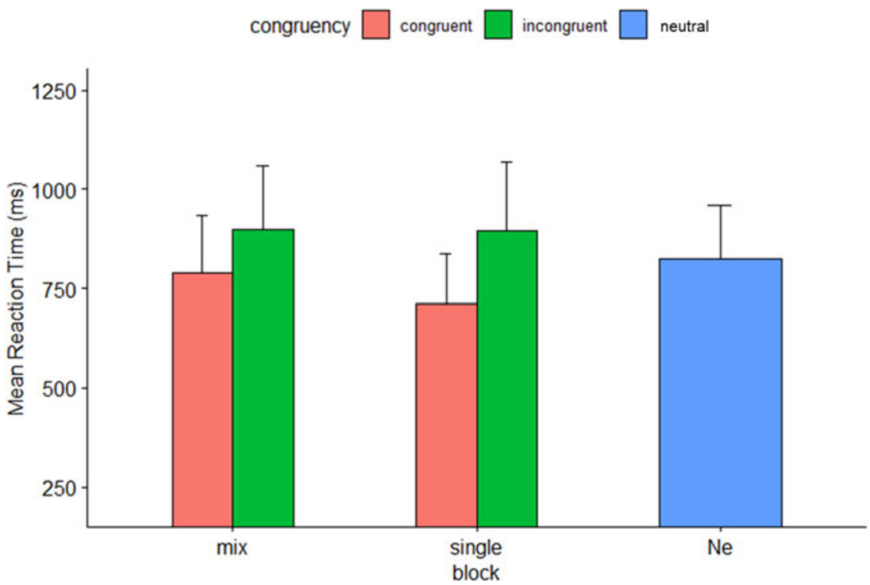

Figure 2. Mean RTs in different trials and blocks in the Verbal Stroop task.

$p < .0001$). Participants' congruent RTs in the mixed block were smaller than their incongruent RTs in the single block, t (39.8) = -7.46, $p < .0001$.

GLMM analysis results revealed that participants' response accuracy differed slightly with congruency ($\beta = -2.87$, SE = 1.35, $z = -2.13$, $p = .03$). However, the pairwise contrasts showed that participants' response accuracy in congruent and incongruent trials was comparable without significant differences across blocks.

RT(s) and response accuracy in the colour-shape switching task

Table 4 shows that participants' mean RTs differed significantly across mix and single-task blocks, F (1, 38.65) = 36.66, $p < .001$. However, no significant effect of trial types on their RTs was found, F (1, 34.41) = 0.20, $p = .66$.

In the mixed block, participants' switch and repeated trials' RTs were comparable (t (39.6) = -0.45, $p = 1.00$), while their single-task trials RTs were significantly smaller than repeated (t (40.0) = 6.05, p <.0001) and switch (t (40.0) = 5.96, p <.0001) trials RTs. Figure 5 illustrates participants' mean RTs for three different trials.

Results showed that participants' response accuracy differed significantly across single and mixed blocks (β =0.89, SE = 0.14, $z = 6.19$, p <.001); while no significant within-block differences were found in the mixed block (β =0.17, SE = 0.11, $z = 1.56$, p =.12). Specifically, participants' response accuracy for repeated trials was smaller than in single-task trials ($\beta = -0.89$, SE = 0.14, $z = -6.19$, $p < .0001$). Similarly, they performed more accurately for single-task trials than switch trials ($\beta = 0.72$, SE = 0.15, $z = 4.96$, p <.0001).

Regression analyses of habitual bilingual language use on verbal and nonverbal cognitive control

Regression analyses were conducted on participants' verbal and nonverbal cognitive control task performance with predictors of

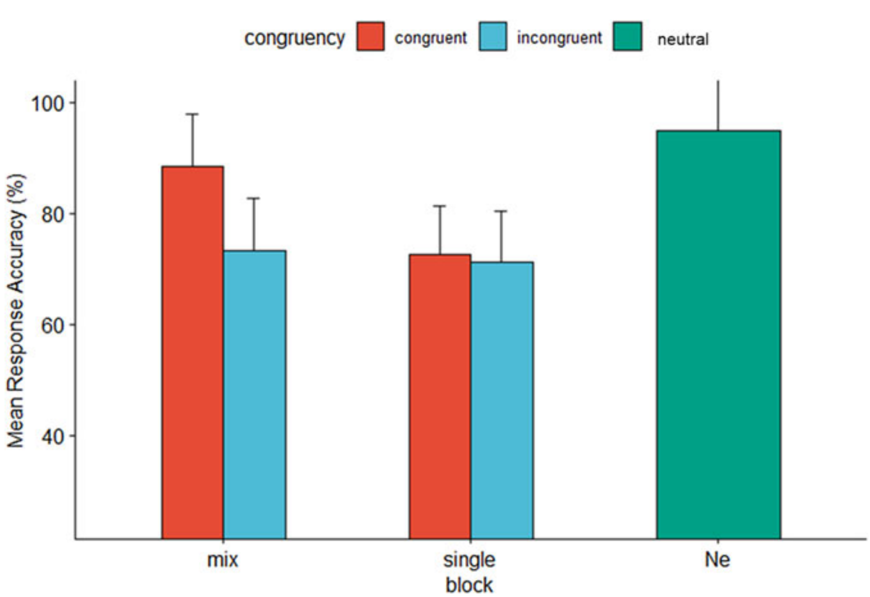

Figure 3. Mean response accuracy differences across mix, single and neutral blocks in the Verbal Stroop task

https://doi.org/10.1017/S1366728924000191 Published online by Cambridge University Press

Table 3. Participants' RTs and response accuracy in the spatial Stroop task

		RTs (ms)		Accuracy (%)	
Block	Trial	Mean	SD	Mean	SD
Neutral (Ne)	neural	465.88	90.93	99.49	1.38
Single	congruent	452.45	76.88	99.49	1.67
	incongruent	557.41	97.09	94.83	13.20
Mix	congruent	498.24	78.16	99.69	1.11
	incongruent	535.59	83.22	92.81	20.22

their different language use experience factors, including L2 proficiency, L2 environment immersion, language entropy in different contexts, habitual language use contexts and language switching habits, and bilingual language production behaviours measured in the two spontaneous language production tasks.

Effects of habitual bilingual language use on participants' verbal Stroop task performance

A random intercept for subject was added to the linear mixed effects model for analysing RTs in the verbal Stroop task. The model reached convergence after a by-subject random slope for the additives of congruency and block was added and reached best-fit convergence after adding z-scored English proficiency and habitual bilingual language use factors, and factors related to participants' performance in the language production tasks (pause frequency ratios, code-switching frequency and pause duration). Models including unstandardised LexTALE test scores were also built, and the results did not differ significantly.

In the subsequent regression analysis, the effects of bilinguals' habitual language use experience (including L2 immersion, self-reported intra/intersentential switching frequency, language entropy in different contexts and L2 proficiency) and participants' spontaneous language production performance (including inter/intrasentential switching frequency in language production tasks and pause ratios in narrative tasks) on their cognitive control performance were considered. Appendix 1 shows the significant factors which influenced participants' RTs in the verbal Stroop task.

The model showed that participants' incongruent RTs in the verbal Stroop task were significantly influenced by their single-language index and intersentential switching index (Hartanto & Yang, 2016). That is, those who habitually use two languages together and code-switching intersententially would perform faster in incongruent trials in this task, showing higher efficiency in verbal inhibitory control.

There was an interactive effect[2] of congruency*L2 environment immersion on participants' verbal inhibitory control performance. It reflected that more intensive L2 environment immersion was associated with bilinguals' overall faster incongruent RTs. Both the interactive effect of block*language entropy in home settings and the effect of block*congruency*language entropy in school settings were found to be positively associated with participants' RTs in this task. That is, participants with lower language entropy in home and school settings had better performance on single-task blocks, especially in the incongruent single-task block, indicating that individuals using English and Chinese separately in home and school settings were proficient in inhibiting and controlling verbal interference in low cognitive monitoring conditions.

Participants' spontaneous language production performance also affected their verbal inhibitory control performance. Participants' pause frequency ratio in bilingual story narration was positively associated with their incongruent verbal Stroop RTs, indicating that fluent bilingual narrators were more efficient in verbal inhibitory control. Besides, the model showed that participants' pause ratios in conversational speech and English narration of Hua Mulan story, as well as their mean pause duration in the English narration of the little match girl story, predicted their incongruent verbal Stroop RTs negatively.

Participants' response accuracy in this task was analysed using a generalized linear mixed effects model. A random intercept for subject was included in the model. No random slopes were included to retain model convergence. The model converged and was significantly improved after adding the factor of L2 environment immersion. The final best-fit model is shown in Table 5.

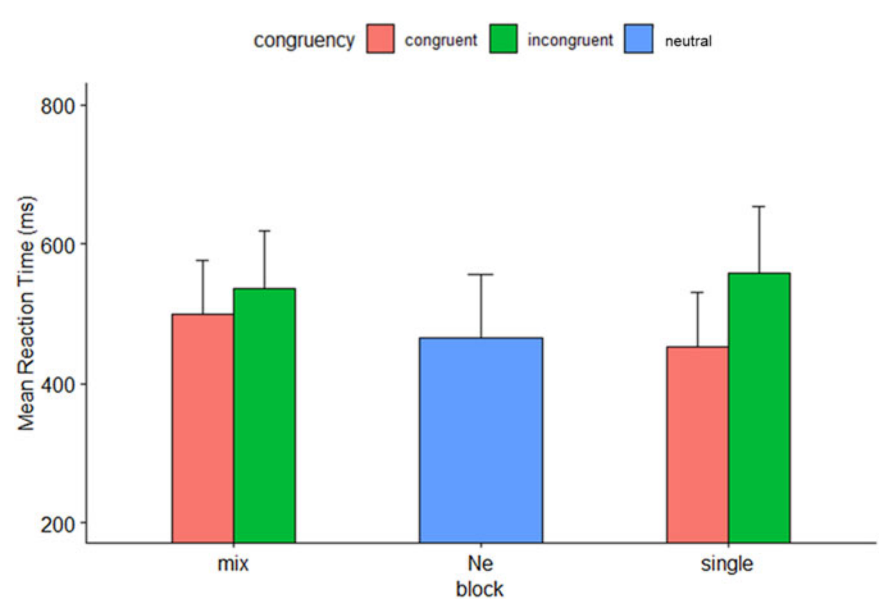

Figure 4. Mean RTs differences across mix, single and neutral blocks in the Spatial Stroop task

https://doi.org/10.1017/S1366728924000191 Published online by Cambridge University Press

Table 4. Participants' RTs and response accuracy performance with switch and mixing costs calculated in the colour-shape switching task

		RTs (ms)		Accuracy (%)	
Block	Trial	Mean	SD	Mean	SD
Mix	Repeated	1218.70	191.00	90.68	11.31
	Switch	1225.70	193.65	91.87	11.60
Single	Single	1051.71	179.94	95.48	10.63
	Switch costs	7.00	85.40	2.27	2.46
	Mixing costs	167.00	181.57	8.43	11.23

The model indicated a negative relationship between bilinguals' intensive L2 environment immersion and their incongruent trials response accuracy. It reflected that bilinguals' accuracy in dealing with verbal interferences might not be significantly improved as their L2 environment immersion increased.

Effects of habitual bilingual language use on participants' spatial Stroop task performance

A random intercept for subject was added to the linear mixed effects model to analyse RTs in the spatial Stroop task, and it reached convergence after a by-subject random slope for the interactives of congruency and block was added. When bilinguals' English proficiency, habitual bilingual language use factors and spontaneous language production performance were fitted into the model, it reached best-fit convergence. The factors which significantly interacted with participants' RTs in the spatial Stroop task were highlighted in the whole best-fit model as shown in Appendix 2.

The model indicated that increasing L2 proficiency and intersentential switching index interconnected with participants' smaller incongruent spatial Stroop RTs, showing their facilitatory effects on bilinguals' nonverbal inhibitory control performance. Moreover, participants with more intensive L2 environment immersion experience tended to perform more efficiently in incongruent trials. Consistent with the RT model in the verbal Stroop task, these findings jointly reflected the modulation of more intensive L2 environment immersion on bilinguals' verbal and nonverbal inhibitory control efficiency.

The model also reflected the effects of participants' language entropy in different contexts on their incongruent RTs. Those participants who had smaller language entropy values in home setting but greater values in settings outside home (i.e., school settings) tended to perform smaller incongruent RTs. Therefore, those single-language context bilingual users, who habitually use their languages in different patterns in distinct contexts, performed more efficiently in the single incongruent spatial Stroop block, which required intensive engagement of proactive inhibitory control.

Interconnecting with participants' performance in the spontaneous language production tasks, the model showed that unbalanced bilinguals with higher fluency in Chinese narrations and frequent intrasentential switching in bilingual story narration had improved nonverbal inhibitory control performance (smaller incongruent RTs). However, their intersentential switching frequency in narrative speech and fluency levels in bilingual conversational speech were not found to affect their nonverbal inhibitory control performance.

To sum up, participants' nonverbal inhibitory control performance was found to be affected by their L2 proficiency, L2 environment immersion and intersentential switching frequency. Moreover, habitual single-language users were found to be more proficient in inhibiting and controlling both verbal and nonverbal interference in low cognitive monitoring conditions.

The interaction between participants' habitual language use and response accuracy in this task was also analysed using a generalized linear mixed effects model. The model included a random intercept for subject. It reached convergence after a by-subject random slope for the interaction of congruency and block was added. The model failed to converge after adding factors related to bilingual habitual language use and participants' spontaneous language production performance. Therefore, participants' habitual language use experience and spontanuous language production performance might not have significant effects on their response accuracy in this task.

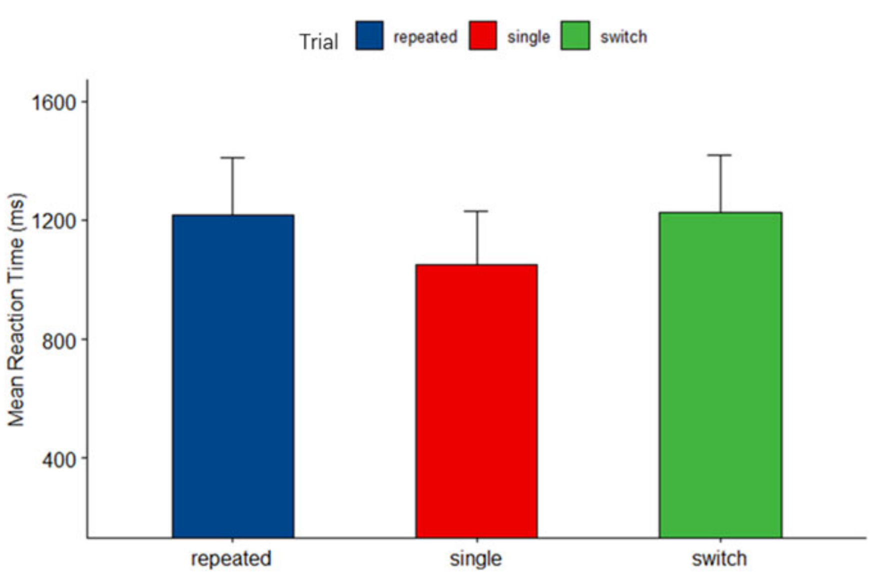

Figure 5. Participants' RTs differences in repeated, switch and single trials in the colour-shape switching task

https://doi.org/10.1017/S1366728924000191 Published online by Cambridge University Press

Table 5. Fixed effects of the general linear mixed effect model for response accuracy in the verbal Stroop task with congruency*block and L2 environment immersion as reference levels. Formula: response accuracy ~ 1 + block *congruency + L2 immersion + (1|subject)

Variable	Estimate	*SE*	z-value	*Pr* (>|z|)
Response accuracy				
(Intercept)	1.53	0.12	13.03	**<.0001**
block (single task)	−0.58	0.08	−7.08	**<.0001**
congruency (incongruent)	−0.54	0.08	−6.65	**<.0001**
block: congruency	0.98	0.16	5.97	**<.0001**
L2 immersion	−0.08	0.03	−2.73	**<.001**

Effects of habitual bilingual language use on participants' colour-shape switching task performance

Participants' RTs switch and mixing costs in the colour-shape switching task were analysed in linear mixed effect models. A random intercept for subject was added to the linear mixed effects model. Since including participants' spontaneous language production performance did not lead to convergence and improvement of the model, these variables were removed from the final model. Factors which significantly affected participants' RTs switch and mixing costs in the colour-shape switching task are highlighted in Appendix 3.

Inconsistent with the study hypothesis, higher L2 proficiency in this model was found to be related to participants' greater RTs switch costs. The result did not reflect the modulation effects of bilinguals' L2 proficiency on their reactive control or shifting efficiency. However, the model indicated a positive relationship between participants' intersentential switching frequency in dual-language contexts and RTs mixing costs. Results further revealed that bilinguals with smaller language entropy in work settings but greater language entropy in school settings tended to have smaller RTs mixing costs. These results indicated that in the habitual single-language context, bilinguals had smaller mixing costs in terms of reaction time, due to an increased efficiency of inhibitory control. Noticeably, the association between habitual single-language use and enhanced proactive inhibition efficiency was more salient among participants with a relatively earlier age of L2 AoA.

Switch and mixing costs for response accuracy were also analysed. Since participants' habitual language use experience and performance in language production tasks did not improve the model, the model only retains participants' language entropy in different contexts as reference levels. A random intercept for subject was added, and the model reached best-fit convergence (see Appendix 4). Consistently, the model also showed that smaller language entropy values in both home and work settings were related to smaller mixing costs, reflecting the significant facilitatory effects of the single-language context on domain-general proactive inhibition efficiency. However, participants' switching costs for response accuracy did not show significant interaction with their habitual language use experience.

Discussion

In this study, effects of individual differences in spontaneous bilingual language production and habitual bilingual language use on bilinguals' domain-general cognitive control performance were investigated.

Relationship between bilinguals' spontaneous language production and their cognitive control efficiency

This study reported that intensive intrasentential switching (i.e., insertion) in narrative speech was significantly correlated with bilinguals' nonverbal inhibitory control performance. In addition, the study indicated a positive interaction between bilinguals' Chinese monolingual narration fluency and their nonverbal inhibitory control efficiency. These results jointly reflected the effects of frequent code-switching on bilinguals' inhibitory control, although bilinguals were relatively unbalanced across two languages, with heightened efficiency in controlling L2 interference in their dominant language production.

Furthermore, the study revealed that higher bilingual narration fluency was associated with bilinguals' enhanced verbal inhibitory control efficiency. This finding suggested that fluent bilingual users, frequently switching between languages, were more efficient in dealing with verbal conflicts, i.e., they suppressed interference from the competing language more efficiently.

However, the positive relationship between intersentential switching frequency in bilingual narrative speech and bilinguals' domain-general inhibitory control efficiency, which is suggested by the ACH and the CPM, was not observed. Similar findings were also revealed in some existing studies (e.g., Hartanto & Yang, 2020; Kałamała et al., 2020), which indicated that the effect of intersentential switching practices is very weak and limited to a complicated interaction between bilingual language experience and inhibitory control development, leading to the findings being hard to interpret (Paap et al., 2021). Furthermore, in line with the previous studies (e.g., Han et al., 2022; Hofweber et al., 2016), this current study showed that bilinguals' intensive intrasentential switching in communication enhanced their proactive inhibition and constant conflict monitoring abilities. The results reflected the multifaceted and complicated interactions between bilingual experience and cognitive control, that the purported models could only be able to explain a limited portion of total variability in the interactions. In addition, although the current results did not fully support the ACH and CPM, it is worthwhile to measure bilinguals' spontaneous language production practices through the naturalistic conversation and narration tasks with relatively high ecological validity.

How does bilinguals' habitual language use experience affect their cognitive control efficiency?

The study revealed that bilinguals' frequent intersentential switching in communication significantly modulated their verbal and nonverbal inhibitory control efficiency. Besides, habitual code-switchers with less engagement in the single-language contexts tended to have enhanced verbal inhibitory control performance. These findings were in line with the predictions of the ACH and the CPM, supporting that higher frequency of intersentential switching modulated bilinguals' efficiency in controlling verbal and nonverbal interference. Such results also reflected that bilinguals' cognitive control processes in language production adapted to their language use patterns (i.e., interchangeably or separately); and frequent language switching in daily communication exercised bilinguals' efficiency, not only in controlling verbal but also nonverbal interference (e.g., Beatty-Martínez et al., 2020; Lai & O'Brien, 2020).

The positive association between bilinguals' L2 environment immersion and their verbal and nonverbal inhibitory control

https://doi.org/10.1017/S1366728924000191 Published online by Cambridge University Press

efficiency highlighted the modulation effects of bilingual language experience on bilinguals' cognitive control. These effects could be attributed to two factors. Firstly, increasing L2 environment immersion could enhance bilinguals' L2 proficiency and consequently increase the extent of activation of L2 and the linguistic competition between L1 and L2 (Green, 1998). Intensive practices of managing co-activated languages and controlling verbal interferences from two proficient languages further exercised bilinguals' efficiency in interference and attentional control (Mishra et al., 2012; Singh & Kar, 2018). Secondly, increasing L2 environment immersion could further strengthen participants' proactive inhibitory control on L1 (e.g., Beatty-Martínez et al., 2020; Woumans et al., 2016), and enhance their proactive control efficiency beyond the language domain (Zhang et al., 2021).

The positive association between participants' language entropy in home and school contexts and their inhibitory control performance revealed in this study was in line with the ACH prediction, i.e., habitually using languages in separate contexts exercised bilinguals' ability in goal maintenance and inhibitory control.

The study showed that participants' shifting performance did not improve as their L2 proficiency increased, which differed from the previously-reported positive relationship (e.g., Tse & Alt Arriba, 2015; J. Wu et al., 2022).It could be that participants in this study were at their ceiling level of shifting; therefore, their task-switching task performance was not sensitive to L2 proficiency development. To address this issue, more studies in future are needed to focus on participants from different age groups, beyond young adults.

However, a negative relationship between bilinguals' single-language contexts engagement and RTs mixing costs was found, while a positive relationship between intersentential switching frequency in dual-language contexts and RTs mixing costs was observed. Additionally, the study revealed a positive association between bilinguals' single-language contexts engagement and response accuracy mixing costs. These findings suggested that habitual single-language context bilinguals exhibit heightened efficiency in proactive control, supporting successful task-set maintenance and consistent monitoring of switching demands. The intensive experience of using two languages separately, based on contextual variations, enhanced their ability in task-set goal maintenance and interference control, particularly in constant conflict monitoring and interference suppression, rather than shifting (Green & Abutalebi, 2013; Han et al., 2022).

In general, the study showed that habitual single-language contexts engagement was effective to modulate bilinguals' sustained inhibition and global cognitive control performance. Moreover, such effects were more salient among bilinguals with earlier L2 AoA; this is in line with existing studies which reported that earlier L2 AoA was associated with bilinguals' better performance in inhibiting prepotent responses (e.g., Soveri et al., 2011; Yow & Li, 2015) and suppressing interference (e.g., Luk et al., 2011). The reason could be that bilinguals with earlier L2 AoA might have more balanced levels in language dominance and consequently produce language switching more often than unbalanced bilinguals (Gollan & Ferreira, 2009; Prior & Gollan, 2011).

Study limitations

Despite the valuable insights gained from our study, it is important to acknowledge its limitations and potential areas for future research.

Firstly, the study involved late Chinese–English bilingual participants who acquired English as a L2 after having Chinese well-acquired. While proficient in both languages, their code-switching predominantly arose from specific contexts or interlocutors. As the researcher is also a Chinese–English bilingual, participants leaned towards using Chinese more frequently. Future studies should involve interlocutors with diverse language backgrounds to examine bilinguals' communication patterns. Furthermore, the study's sample size (N = 41) may limit insights into how various language experiences influence bilinguals' spontaneous language production and cognitive control.

Secondly, employing conversational and narrative tasks to measure bilinguals' spontaneous language production is an innovative but underutilized paradigm in relevant research. More studies adopting this approach are expected. Future research should further differentiate code-switching types in spontaneous language production and refine approaches to capture bilinguals' language experience dynamically.

Thirdly, in this study, the order of the three cognitive control tasks was randomized among participants. Nevertheless, these tasks consistently followed the language production tasks. This order aimed to mitigate withdrawal rates during online data collection amid the pandemic challenges. However, it could potentially divulge the study purpose early, influencing participants' subsequent task performance. For future studies, fully counterbalancing the order of language and cognitive tasks across different participants is recommended.

Conclusion

In conclusion, this study highlighted the significant influences of bilingual experience on spontaneous language production and cognitive control in language processing. It partially supported the predictions of the ACH and the CPM that frequent language switching in dual-language contexts can strengthen bilinguals' domain-general inhibitory control. In contrast, bilinguals habituated to single-language contexts were expert in goal maintenance, sustained conflict monitoring and control rather than task disengagement and engagement.

The study also revealed the positive association between bilinguals' frequent code- switching in communication and better inhibitory control performance. Frequent and fluent code-switchers tended to show enhanced verbal and nonverbal inhibitory control abilities. Although the expected interconnection between code-switching frequency and bilinguals' shifting abilities was not found, the study stressed that code-switching practices, as the significant factor in bilingual experience, can significantly modulate bilinguals' behaviours regarding language control and domain-general inhibitory control.

The study also explored the interactive effects of L2 AoA, language proficiency, and L2 environment immersion on bilinguals' language and cognitive control performance. Results addressed the significant interconnections among these factors and revealed the different magnitudes of effects on bilinguals' cognitive control derived from their individual differences in language use experience. This study also suggested the necessity to take individual differences in bilingualism development into account when characterising the effects of bilingual experience on cognitive control.

Supplementary Material. For supplementary material accompanying this paper, visit https://doi.org/10.1017/S1366728924000191

Data availability. The data that support the findings of this study are openly available via the Open Science Framework in *Bilingual language experience in spontaneous language production and cognitive control* at http://doi.org/10.17605/OSF.IO/7XHDN.

Acknowledgements. Many thanks to all participants that took part in the study during the pandemic and enabled this research to be possible. Our thanks to the anonymous reviewers for their valuable comments that improved our work greatly.

Notes

[1] To contrast with intersentential switching, the term "intrasentential switching" is used to denote the specific code-switching type, insertion, throughout the study.

[2] In the interactive model, the interactive effects of congruency and L2 environment immersion (congruency:Yrs_in_EN), and the interactive effects of block and language entropy (e.g., block:home_entropy) were included. The interactive effects between congruency and block on participants' cognitive control tasks were also considered (i.e., congruency:block).

References

Acheson, D. J., Ganushchak, L. Y., Christoffels, I. K., & Hagoort, P. (2012). Conflict monitoring in speech production: Physiological evidence from bilingual picture naming. *Brain and Language*, *123*(2), 131–136.

Barac, R., & Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child development*, *83*(2), 413-422.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1).

Beatty-Martínez, A. L., & Dussias, P. E. (2017). Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language*, *95*, 173-189.

Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(6), 1022–1047.

Blanco-Elorrieta, E., & Pylkkänen, L. (2017). Bilingual Language Switching in the Laboratory versus in the Wild: The Spatiotemporal Dynamics of Adaptive Language Control. *The Journal of Neuroscience*, *37*(37), 9022–9036.

Blumenfeld, H. K., & Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition*, *118*(2), 245-257.

Blumenfeld, H. K., & Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology*, *25*(5), 547-567.

Boersma, P., & Weenink, D. (2020). Praat: Doing phonetics by computer [Computer program] (Version 6.1.24).

Bonfieni, M., Branigan, H. P., Pickering, M. J., & Sorace, A. (2019). Cognitive control in bilinguals: Effects of language experience and individual variability. *Bilingualism: Language and Cognition*, 1–12.

Choo, K. K., Keat, W. H., & Price, J. (2021). *Modulatory Effects of Language Context on Executive Functioning in Multilinguals* [Preprint]. PsyArXiv.

DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences*, *116*(15), 7565–7574.

Gardner-Chloros, P., McEntee-Atalianis, L., & Paraskeva, M. (2013). Code-switching and pausing: An interdisciplinary study. *International Journal of Multilingualism*, *10*(1), 1–26.

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*(3), 640–665.

Gollan, T. H., Kleinman, D., & Wierenga, C. E. (2014). What's Easier: Doing What You Want, or Being Told What to Do? Cued versus Voluntary Language and Task Switching. *Journal of Experimental Psychology. General*, *143*(6), 2167–2195.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, *1*(2), 67-81.

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

Green, D. W., & Li, W. (2014). A control process model of code-switching. *Language, Cognition and Neuroscience*, *29*(4), 499–511.

Grosjean, F. (2013). Bilingualism: A short introduction. *The psycholinguistics of bilingualism*, *2*(5).

Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, *23*(2), 283–294.

Han, X., Li, W., & Filippi, R. (2022). The effects of habitual code-switching in bilingual language production on cognitive control. *Bilingualism: Language and Cognition*, 1–21.

Hartanto, A., & Yang, H. (2016). Disparate bilingual experiences modulate task-switching advantages: A diffusion-model analysis of the effects of interactional context on switch costs. *Cognition*, *150*, 10–19.

Hartanto, A., & Yang, H. (2020). The role of bilingual interactional contexts in predicting interindividual variability in executive functions: A latent variable analysis. *Journal of Experimental Psychology: General*, *149*(4), 609–633.

Heidlmayr, K., Moutier, S., Hemforth, B., Courtin, C., Tanzmeister, R., & Isel, F. (2014). Successive bilingualism and executive functions: The effect of second language use on inhibitory control in a behavioural Stroop Colour Word task. *Bilingualism: Language and Cognition*, *17*(3), 630–645.

Hlavac, J. (2011). Hesitation and monitoring phenomena in bilingual speech: A consequence of code-switching or a strategy to facilitate its incorporation?. *Journal of Pragmatics*, *43*(15), 3793-3806.

Hofweber, J., Marinis, T., & Treffers-Daller, J. (2016). Effects of dense code-switching on executive control. *Linguistic Approaches to Bilingualism*, *6*(5), 648–668.

Kałamała, P., Szewczyk, J., Chuderski, A., Senderecka, M., & Wodniecka, Z. (2020). Patterns of bilingual language use and response inhibition: A test of the adaptive control hypothesis. *Cognition*, *204*, 104373.

Kheder, S., & Kaan, E. (2021). Cognitive control in bilinguals: Proficiency and code-switching both matter. *Cognition*, *209*, 104575.

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*(5), 700–714.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). **lmerTest** Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13).

Lai, G., & O'Brien, B. A. (2020). Examining Language Switching and Cognitive Control Through the Adaptive Control Hypothesis. *Frontiers in Psychology*, *11*.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343.

Lloyd-Smith, A., Einfeldt, M., & Kupisch, T. (2020). Italian-German bilinguals: The effects of heritage language use on accent in early-acquired languages. *International Journal of Bilingualism*, *24*(2), 289-304.

Luk, G., De Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism: Language and Cognition*, *14*(04), 588–595.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967.

Meuter, R. F. I., & Allport, A. (1999). Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection. *Journal of Memory and Language*, *40*(1), 25–40. https://doi.org/10.1006/jmla.1998.2602

Mishra, R. K., Hilchey, M. D., Singh, N., & Klein, R. M. (2012). On the time course of exogenous cueing effects in bilinguals: higher proficiency in a second language is associated with more rapid endogenous disengagement. The *Quarterly journal of experimental psychology*, *65*(8), 1502-1510.

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

Navracsics, J. (2004). The question of control in bilingual speech production in different language modes. *Grazer Linguistische Studien*, (*62*).

Ng, G., & Yang, H. (2021). Code-Switching patterns differentially shape cognitive control: Testing the predictions of the adaptive control hypothesis. *Bilingualism: Language and Cognition*, 1–15.

Ooi, S. H., Goh, W. D., Sorace, A., & Bak, T. H. (2018). From Bilingualism to Bilingualisms: Bilingual experience in Edinburgh and Singapore affects attentional control differently. *Bilingualism: Language and Cognition*, *21* (4), 867–879.

Paap, K. R., Mason, L., & Anders-Jefferson, R. (2021). Predictions about the cognitive consequences of language switching on executive functioning inspired by the adaptive control hypothesis fail more often than Not. *Brain Sciences*, *11*(9), 1217.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Poplack, S. (1980). Sometimes i'll start a sentence in spanish y termino en espanol: toward a typology of code-switching1.

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: JINS*, *17*(4), 682–691.

Révész, A., Ekiert, M., & Torgersen, E. N. (2016). The effects of complexity, accuracy, and fluency on communicative adequacy in oral task performance. *Applied linguistics*, *37*(6), 828-848.

Rodriguez-Fornells, A., Krämer, U. M., Lorenzo-Seva, U., Festman, J., & Münte, T. F. (2012). Self-Assessment of Individual Differences in Language Switching. *Frontiers in Psychology*, *2*.

R Studio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Schilperoord, J. (2002). On the cognitive status of pauses in discourse production. *Contemporary tools and techniques for studying writing*, 61-87.

Singh, J. P., & Kar, B. R. (2018). Effect of language proficiency on proactive occulo-motor control among bilinguals. *PLOS ONE*, *13*(12), e0207904.

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*.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, *18*(6), 643.

Surrain, S., & Luk, G. (2019). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*, *22*(2), 401–415.

Toribio, A. J. (2001). On the emergence of bilingual code-switching competence. *Bilingualism: language and cognition*, *4*(3), 203-231.

Treffers-Daller, J. (2009). Code-switching and transfer: An exploration of similarities and differences. In B. E. Bullock & A. J. Toribio (Eds.), *The Cambridge Handbook of Linguistic Code-switching* (pp. 58–74). Cambridge University Press.

Tse, C.-S., & Alt Arriba, J. (2015). Local and Global Task Switching Costs in Bilinguals Who Vary in Second Language Proficiency. *The American Journal of Psychology*, *128*(1), 89–106.

Woumans, E., Surmont, J., Struys, E., & Duyck, W. (2016). The Longitudinal Effect of Bilingual Immersion Schooling on Cognitive Control and Intelligence*. *Language Learning*, *66*(S2), 76–91.

Wu, Y. J., & Thierry, G. (2013). Fast Modulation of Executive Function by Language Context in Bilinguals. *Journal of Neuroscience*, *33*(33), 13533–13537.

Wu, J., Wang, Q., Wu, X., Ji, Y., & Yan, G. (2022). Second Language Proficiency Modulates the Dependency of Bilingual Language Control on Domain-General Cognitive Control. *Frontiers in Psychology*, *13*.

Yim, O., & Bialystok, E. (2012). Degree of conversational code-switching enhances verbal task switching in Cantonese–English bilinguals. *Bilingualism: Language and Cognition*, *15*(4), 873-883.

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*.

Zantout, M. (2019). *Investigating multilingual adolescents' cognitive and codeswitching performance: an exploratory application of language control models*. Ph.D. dissertation, University of Oxford.

Zhang, H., Diaz, M. T., Guo, T., & Kroll, J. F. (2021). Language immersion and language training: Two paths to enhanced language regulation and cognitive control. *Brain and Language*, *223*, 105043.