CAMBRIDGE UNIVERSITY PRESS

Modality-general and modality-specific bilingual control mechanisms in spoken and written productions

Tingting Yang[1], Zhenguang G. Cai[2], Weihao Lin[1] and Ruiming Wang[1] (iD)

[1]Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents, Ministry of Education, & Centre for Studies of Psychological Application, School of Psychology, South China Normal University, Guangzhou, China and [2]Department of Linguistics and Modern Languages & Brain and Mind Institute, The Chinese University of Hong Kong, Hong Kong SAR, China

Abstract

Do bilinguals have similar bilingual control mechanisms in speaking and writing? The present study investigated the patterns of switch costs (reflecting reactive language control) and mixing costs (reflecting proactive language control) between Chinese (L1) and English (L2) in spoken and written productions and whether these patterns could be modulated by response-stimulus intervals (RSIs). In two experiments, unbalanced Chinese–English bilinguals completed a cued language switching task in spoken naming (Experiment 1) and written naming (Experiment 2), respectively. The results revealed asymmetrical switch costs (i.e., the larger cost in L1 than in L2) in spoken and written productions in the short RSI condition. However, there were asymmetrical mixing costs in spoken production and symmetrical mixing costs in written production both in the short and long RSIs. These findings suggest that for spoken and written productions, reactive language control operates in similar mechanisms, while proactive language control operates in specific mechanisms.

1. Introduction

A growing body of research suggests that both the target and non-target languages are activated in bilingual production (Costa & Santesteban, 2004; Declerck et al., 2012; Philipp et al., 2007). The activation of non-target language can interfere with the selection and access to the target language (Green, 1998; Meuter & Allport, 1999), requiring bilinguals to recruit a language control mechanism to detect and resolve this cross-language interference.

One commonly used method for investigating bilingual language control is the cued language switching paradigm, where bilinguals name pictures or digits in their native language (L1) or second language (L2) based on a given language cue (e.g., a flag or color frame) in single language blocks and mixed language blocks. In mixed language blocks, the naming language of two sequential trials can be the same (repeat trials) or different (switch trials). Two typical findings have been repeatedly observed. First, there is a language switch cost: bilinguals perform slower and more erroneous for switch trials than for repeat trials in a mixed language block (Abutalebi & Green, 2007; Bialystok et al., 2006). Second, there is a language mixing cost: bilinguals perform worse for repeat trials in the mixed language blocks than for trials in the single language blocks (Christoffels et al., 2007; Declerck, 2020; Jylkkä et al., 2018; Prior & Gollan, 2011).

It is worth noting that the majority of research on bilingual language control has focused on spoken production, with little attention given to written production. However, previous monolingual studies have shown that conceptual and lexical-semantic processes are shared between spoken and written productions, while phonological/orthographic encoding and motor execution processes are separate (Breining & Rapp, 2019; Muylle et al., 2022; Perret et al., 2014). With the increasing prevalence of multilingualism and multiculturalism in today's society, it is becoming increasingly common for people to speak or write in more than one language. Additionally, many second language learners rely on written input (e.g., from print; Tainturier, 2019). Therefore, it is essential to explore how the bilingual control mechanism may differ between spoken and written productions.

Bilingual language control has been typically divided into two types of control processes: reactive language control and proactive language control (Declerck, 2020; Ma et al., 2016; Peeters & Dijkstra, 2018). Reactive language control is engaged when the non-target language interferes with the selection of target language words and is more transient and local (trial-by-trial) in essence. Indicators of reactive language control include the language switch

cost and the n-2 language repetition cost (Declerck & Koch, 2023; Gade et al., 2021). The n-2 language repetition cost is measured when a mixed language block has three naming languages. For example, using ABC to represent three different languages respectively, the n-2 language repetition costs refer to the worse performance in n-2 language repetition trials (i.e., trial A at the end of the ABA sequence) than in n-2 language switch trials (i.e., trial A at the end of the CBA sequence).

Proactive language control, on the other hand, is reflected in anticipation of producing a target language and the proactive inhibition of the non-target language. It is more contextually induced, sustained, and global in nature. Indicators of proactive language control include the language mixing cost (as discussed above), the reversed language dominance effect, and the blocked language order effect. The reversed language dominance effect refers to performance being worse when naming in the dominant language than in the non-dominant language within mixed language blocks, while the blocked language order effect refers to the worsened performance in a single language block following another language block than in a single language block alone. Therefore, while reactive language control addresses trial-by-trial cross-language interference, proactive language control operates as a preventative mechanism at a more global level (for reviews, see Declerck, 2020; Declerck & Philipp, 2015).

Language switch costs are often asymmetrical in unbalanced bilinguals (i.e., the switch costs are greater in L1 than in L2; Gollan et al., 2014; Ma et al., 2016; Meuter & Allport, 1999) and symmetrical in balanced bilinguals (i.e., the switch costs are comparable between L1 and L2; Costa & Santesteban, 2004; Costa et al., 2006). According to the inhibitory control model, bilinguals control their languages by inhibiting the activation level of the non-target language while accessing lexical representations in the target language. Asymmetrical switch costs suggest that unbalanced bilinguals inhibit their L1 during L2 processing to a greater extent than inhibiting their L2 during L1 processing. Furthermore, some studies have found that language mixing costs are also asymmetrical across languages for unbalanced bilinguals, with larger mixing costs in L1 than in L2 (e.g., Jylkkä et al., 2018; Ma et al., 2016; Mosca & de Bot, 2017; Peeters & Dijkstra, 2018). It has been proposed that, for unbalanced bilinguals, the more dominant L1 might exert stronger sustained or global inhibition for the dominant L1 than the non-dominant L2 in the bilingual context (Ma et al., 2016; Peeters & Dijkstra, 2018), while the baseline activation of L1 is generally higher than that of L2 in the single language context (Broos et al., 2021; Hanulová et al., 2011). Therefore, unbalanced bilinguals may engage in stronger proactive language control in L1 than in L2.

There is evidence that several task-related parameters can impact the bilingual control mechanisms in spoken production (Bobb & Wodniecka, 2013; Fink & Goldrick, 2015; Ma et al., 2016; Verhoef et al., 2009). One critical parameter is the cue-to-stimulus interval (CSI), which is the time between the cue onset and the stimulus onset. Another critical parameter is the response-stimulus interval (RSI), which is the time between the onset of naming the stimulus in a trial and the onset of the stimulus in the next trial (Bobb & Wodniecka, 2013; Mosca & Clahsen, 2016).

Previous studies have shown that the CSI can affect the magnitude of language switch costs, with switch costs typically decreasing as the CSI increases (Costa & Santesteban, 2004; Verhoef et al., 2009). However, whether the CSI also impacts

the degree of asymmetry in switch costs is unclear: some studies found that the CSI can modulate the asymmetry of switch costs (Liu et al., 2019; Verhoef et al., 2009), while others did not (Fink & Goldrick, 2015; Ma et al., 2016; Philipp et al., 2007).

RSI may also play a crucial role in the patterns of asymmetrical switch costs (Declerck et al., 2012). Philipp et al. (2007) observed asymmetrical switch costs with both short CSI (100 ms) and long CSI (1000 ms) when the RSI was fixed (1100 ms). However, Verhoef et al. (2009) showed asymmetrical switch costs in short CSI (750 ms) and symmetrical switch costs in long CSI (1500 ms), but with variable RSIs in the short CSI (in which RSI range from 2250 ms to 3050 ms) and long CSI (in which RSI range from 3000 ms to 3800 ms) in their study. Another interpretation of these findings is that a short RSI leads to asymmetrical switch costs, while a long RSI leads to symmetrical switch costs.

Ma et al. (2016) conducted the first study to independently manipulate CSI and RSI to investigate the effect on switch costs and mixing costs in bilingual production. In one experiment, when RSI was constant, they found that as CSI increased, the switch costs and mixing costs decreased. Additionally, they discovered that CSI affected the asymmetry of mixing costs but not that of switch costs. In another experiment, when setting the cue and target to be presented simultaneously (i.e., without a CSI), they observed that switch costs decreased as RSI decreased, but mixing costs remained unchanged. They also found that RSI affects the asymmetry of switch costs and mixing costs, with greater asymmetry in switch costs at shorter RSI and greater asymmetry in mixing costs at longer RSI. Both CSI and RSI are important parameters for manipulating reactive language control (measured by switch costs) and proactive language control (measured by mixing costs) in L1 and L2. Notably, previous task switching research has shown that switch costs can be reduced by lengthening RSI or CSI (Grange & Cross, 2015; Horoufchin et al., 2011; Koch & Allport, 2006). In fact, CSI reflects the active preparation of current task requirements. In the absence of CSI, RSI reflects the dissipation of activation of the previous task (i.e., passive decay; Kiesel et al., 2010; Koch & Allport, 2006).

Wong and Maurer (2021) used a sequence-based language switching paradigm and found that switch costs did not differ between spoken and written productions. They suggest that some aspects of bilingual language control may be similar across these modalities. It is worth noting that their paradigm made use of predictable sequences of language switches (e.g., L1-L1-L2-L2-L1-L1-L2-L2…) without a language cue, and RSI was fixed at 4 seconds for both production modalities. The RSI in the study might thus be quite long for spoken production, which may explain the lack of a difference in switch costs between production modalities. In some monolingual studies, RSI is set longer for written production than spoken production (e.g., Damian et al., 2011; Qu et al., 2021). Handwriting a word takes longer time than speaking a word (Gould & Boies, 1978). When RSI is kept constant, the remaining time after response completion is longer for spoken production than for written production. Therefore, it is necessary to set multiple RSI values to fully understand the effects of it on bilingual control mechanisms in spoken and written productions.

The present study aims to explore the similarities and differences in language control (reactive language control and proactive language control) between spoken and written productions by manipulating RSI lengths. Two experiments were conducted to achieve this goal. In Experiment 1, we aimed to investigate

whether the patterns of switch costs and mixing costs between Chinese (L1) and English (L2) are modulated by RSI lengths in spoken naming. In this experiment, unbalanced Chinese–English bilinguals performed cued language switching in spoken picture naming. The range of RSI length used was longer than in a previous study (Ma et al., 2016), with 2000 ms (short RSI) and 3500 ms (long RSI) being used. This was done to ensure that participants had enough time to write the whole word for comparison between spoken and written tasks. In Experiment 2, the same design was used for written naming. This experiment aimed to explore whether the patterns of switch and mixing costs can be affected by the varying RSI lengths in written production.

By combining the results of both experiments, the study aims to summarize the similarities and differences in language control between spoken and written productions. Suppose similar patterns of switch costs or mixing costs are found for speaking and writing, and these costs are similarly sensitive to the RSI manipulation. In that case, it is expected that this kind of language control operates in similar mechanisms across the two production modalities, i.e., modality-general. Conversely, if the results show different patterns, it is considered that this kind of language control operates in specific mechanisms between spoken and written productions, i.e., modality-specific. The study aims to infer the modality-general and modality-specific bilingual control mechanisms between speaking and writing.

2. Experiment 1

2.1. Participants

Fifty-six Chinese–English bilingual participants[1] from the South China Normal University were recruited for the present experiment (39 females; Mage = 20.77, SDage = 1.69). All participants were right-handed, had normal or corrected-to-normal vision, and had no reported history of neurological impairments or language disorders. They all signed a written informed consent form. This experiment was approved by the local authority approved. Participants were randomly assigned to the different RSI conditions for a cued language switching task in spoken naming, with 28 participants in the short RSI condition and 28 participants in the long RSI condition.

All participants were non-English major students and self-reported their proficiency in listening, speaking, reading, and writing on a 7-point scale (1 = not fluent at all, 7 = very fluent) for both Chinese and English. They also completed a lexTALE test (Lemhöfer & Broersma, 2012) to assess their English vocabulary size. An independent sample t-test showed that there was no difference between the two groups of participants in age, proficiency scores for Chinese and English in the four skills, L2 age of acquisition (AoA), and lexTALE test scores (all $ps > .1$). See Table 1 for detailed participants' information. In addition, paired samples t-tests revealed that proficiency scores for Chinese were significantly higher than for English in listening ($t = 16.12$, $p < .001$), speaking ($t = 18.35$, $p < .001$), reading ($t = 11.65$, $p < .001$), and writing ($t = 12.38$, $p < .001$), indicating that the participants were unbalanced bilinguals with a dominant L1.

We also used G*power (Version 3.1.9; Faul et al., 2007) to estimate the (post hoc) power of the experiment, treating the interaction effects as difference scores. With $N = 28$ for each group and an expected effect size of $d_z = 0.8$, this experiment achieved a power of 83.6%. Therefore, the sample size in this experiment was acceptable.

2.2. Materials

In this study, ten black-and-white pictures (line-drawings) were selected from the database of Zhang and Yang (2003). Of these, eight were used as experimental items, and two were used as filler items. Each picture had a pair of Chinese and English names (see Appendix A). The eight experimental items all had monosyllabic Chinese names (e.g., 月), with a mean stroke number of 5.1 (range from 3 to 7); their English names (e.g., moon) had a mean letter number of 4.1 (range from 3 to 6). The same pictures were presented in L1-single, L2-single, and mixed-language contexts.

2.3. Task and procedure

Participants first familiarized themselves with the pictures and their names by seeing each picture together with its printed Chinese and English names. The main experiment was run on E-Prime 2.0 on a desktop. A trial began with a fixation cross (+) in the middle of the screen for 500 ms, followed by a picture framed inside a red or green square. Participants were to name the picture as quickly and accurately as possible in the language indicated by the color of the frame (i.e., Chinese when it was red and English when it was green for half of the participants, and the reverse for the other half). The picture disappeared once participants initiated a spoken response or did not respond within 3 seconds.

We digitally recorded naming latencies, measured from picture onset to voice onset, using a voice key connected to the computer via a Chronos Response Box. The interval from response to the next fixation onset was 1500 ms or 3000 ms; therefore, the interval from response to the next stimuli onset was 2000 ms and 3500 ms (i.e., the short and long RSI conditions, respectively; see Figure 1).

The experiment first conducted two single-language blocks (one for Chinese and one for English, as indicated by the color cue, with the language order counterbalanced across participants) followed by four mixed-language blocks. Each single-language block had 48 experimental trials. Each mixed-language block included one filler trial (the first trial, as it was neither the switch trial nor the repeat trial) and 48 experimental trials. The eight experimental pictures were presented three times for each language in each block. Each single-language block was preceded by 10 practice trials; and the first mixed block was preceded by 20 practice trials. Participants took a short break between the practice trials and the experimental block and after every experimental block. There were 48 trials for each trial type (L1 single, L2 single, L1 repeat, L2 repeat, L1 switch, and L2 switch), with every experimental picture (i.e., the eight experimental items) presented 6 times for each trial type. The whole experiment was video-recorded using the video recording software EV Capture, which is in order to check the speech errors manually. The experiment lasted for about 35 minutes in the short RSI condition and about 45 minutes in the long RSI condition.

2.4. Results

We first removed the trials that were not appropriately recorded due to the failure of voice-key triggering (0.95% of all trials). Next, incorrect trials where a wrong word or a wrong language was used were discarded, and in cases where a wrong language was used, the subsequent trials were also discarded (3.55% of all trials). Then, naming latencies were trimmed by discarding trials

Table 1. Participant characteristics in Experiment 1 (with SD in parentheses)

		Short RSI condition	Long RSI condition	t	p
	N (Female)	28 (19)	28 (20)		
	Age	21.00 (1.68)	20.54 (1.71)	1.03	.310
Chinese	Listening	6.29 (0.81)	6.36 (0.78)	-0.34	.738
(L1)	Speaking	5.82 (0.98)	6.14 (0.93)	-1.26	.215
	Reading	6.04 (0.74)	6.25 (0.80)	-1.04	.304
	Writing	5.29 (0.94)	5.43 (1.14)	-0.51	.610
English	Listening	3.68 (1.06)	3.29 (1.08)	1.37	.175
(L2)	Speaking	3.43 (0.79)	3.11 (0.96)	1.37	.176
	Reading	4.64 (0.73)	4.32 (1.12)	1.27	.210
	Writing	3.68 (1.09)	3.64 (1.16)	0.12	.906
	L2_AoA	7.93 (2.11)	8.04 (2.38)	-0.18	.859
	LexTALE	58.08 (8.65)	57.77 (8.20)	0.14	.890

with a latency below 300 ms or above 2500 ms (0.39% of all trials). Lastly, latencies of 2.5 standard deviations above or below the participants' mean were removed (2.96% of all trials) to obtain the final retained latencies data. The log transformation was applied to latencies data to reduce its skewness (pre-transformation skewness = 0.90 and post-transformation skewness = 0.18). The error trials for accuracy analyses included using wrong words and language. Table 2 shows the mean latencies for Experiment 1.

The data were analyzed to determine switch costs and mixing costs separately. Switch costs were determined by comparing log-transformed latencies (hereafter referred to as log latencies) and accuracy between switch trials and repeat trials. Mixing costs were determined by comparing log latencies and accuracy between repeat trials in the mixed-language context and trials in the single-language context. Linear mixed-effects models (for log latencies) and logistic mixed-effects models (for correct vs

incorrect responses) were used (lmerTest package; Kuznetsova et al., 2017) in the statistical software R (version 4.2.2). Deviation contrasts (-0.5 and 0.5) were used for all fixed effects to determine the main effects and interactions. In the models, we included RSI (short = -0.5, long = 0.5), Language (L1 = -0.5, L2 = 0.5), Trial type (switch costs: repeat = -0.5, switch = 0.5; mixing costs: single = -0.5, repeat = 0.5), and their interactions, as fixed effects. The full model used the maximal random effect structure, including all random intercepts and all random slopes for participants and items (Barr et al., 2013). If the full model did not converge, then we adopted a backward-stepping procedure to sequentially reduce the random slopes until the model could be fitted. In addition, if several models with various random structures can be fitted, the optimal model was selected according to AIC criteria. Post-hoc multiple comparisons used the emmeans package (Lenth, 2022). The experimental data and analytical scripts are publicly available at https://osf.io/yhs4m/.
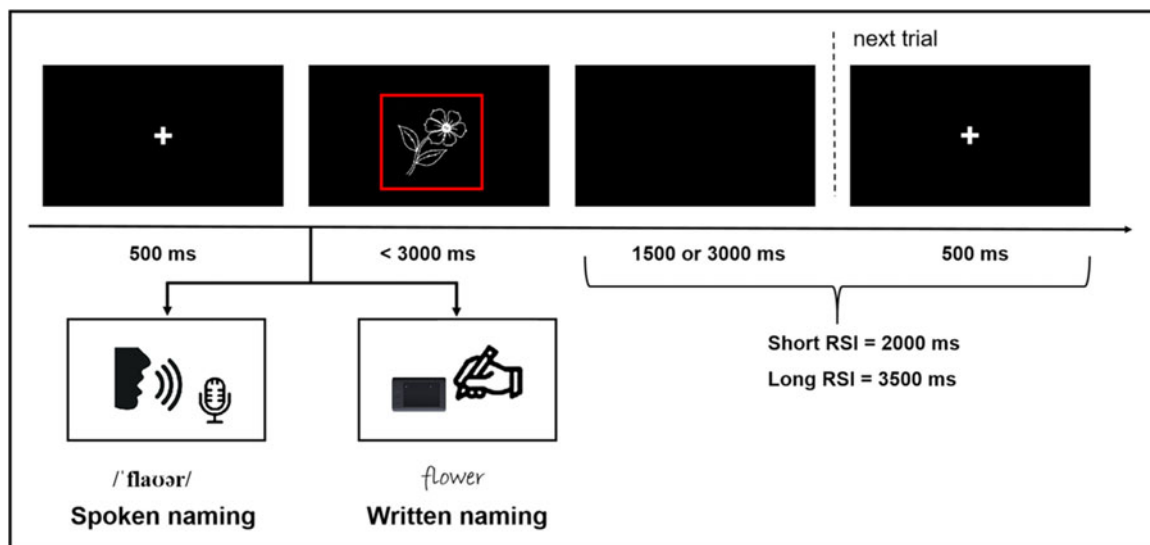

Figure 1. Example of a trial sequence in Experiment 1 (spoken naming) and Experiment 2 (written naming). The spoken naming latency was the time that the voice onset took contact with the voice key after the picture onset. The written naming latency was the time that the pen point took to initial contact with the graphic tablet after picture onset.

Table 2. Mean naming latency (with SD in parentheses) and accuracy (with SD in parentheses) as a function of trial type across two response-stimulus intervals (RSIs) in Experiment 1

		Switch	Repeat	Single
Latency (ms)				
Short RSI	L1	871 (191)	783 (185)	627 (135)
	L2	833 (185)	769 (167)	671 (119)
Long RSI	L1	892 (216)	837 (209)	648 (123)
	L2	884 (201)	821 (195)	694 (130)
Accuracy (%)				
Short RSI	L1	95.9 (4.0)	98.1 (2.9)	100.0 (0.0)
	L2	96.9 (2.9)	98.9 (1.2)	99.9 (0.5)
Long RSI	L1	95.7 (4.1)	97.8 (2.7)	99.9 (0.5)
	L2	97.7 (2.5)	99.2 (1.5)	99.9 (0.4)

Switch costs

The analysis of log latencies revealed a marginally significant main effect of Language ($\beta = 0.020$, $SE = 0.010$, $t = 1.94$, $p = .077$), with longer latency in L1 than in L2 ($M = 845 \pm 205$ ms and $M = 826 \pm 191$ ms, respectively). In addition, the main effect of Trial type reached significance ($\beta = 0.084$, $SE = 0.009$, $t = 9.73$, $p < .001$), with longer latency in switch trials compared to repeat trials ($M = 870 \pm 199$ ms and $M = 803 \pm 191$ ms, respectively). Notably, we found a significant three-way interaction among RSI, Language, and Trial type ($\beta = 0.039$, $SE = 0.016$, $t = 2.46$, $p = .014$). Further analysis showed that the switch cost was significantly greater in L1 than L2 in short RSI ($M = 87$ ms and $M = 66$ ms, respectively; $z = 2.43$, $p = .015$), while it was comparable between L1 and L2 in long RSI ($M = 60$ ms and $M = 69$ ms, respectively; $z = -1.05$, $p = .296$). These results indicate an asymmetry in switch costs in the short RSI condition and a symmetry in switch costs in the long RSI condition (see Figure 2). In addition, following the analytical logic in Verhoef et al. (2009), we further analyzed this three-way interaction in terms of the RSI effect (i.e., faster latency for short RSI compared to long RSI). Further planned comparisons revealed that there was a marginally significant RSI effect in L1 repeat trials ($z = 1.79$, $p = .073$), L2 switch trials ($z = 1.94$, $p = .053$), and L2 repeat trials ($z = 1.85$, $p = .064$), but not in L1 switch trials ($z = 0.76$, $p = .450$).

The analysis of accuracy revealed the significant main effects of Language ($\beta = 0.479$, $SE = 0.183$, $z = 2.62$, $p = .009$) and Trial type ($\beta = 0.890$, $SE = 0.141$, $z = 6.32$, $p < .001$), with higher accuracy in L2 than in L1 ($M = 98.2\% \pm 1.6\%$ and $96.9\% \pm 2.9\%$, respectively) and higher accuracy in repeat trials compared to switch trials ($M = 98.5\% \pm 1.7\%$ and $M = 96.6\% \pm 2.8\%$, respectively). No other main effects or interactions were significant ($ps > .1$).

Mixing costs

The analysis of log latencies revealed a significant main effect of RSI ($\beta = 0.048$, $SE = 0.029$, $t = 1.69$, $p = .097$), with shorter latency in short RSI than in long RSI ($M = 710 \pm 166$ ms and $M = 747 \pm 186$ ms, respectively). In addition, the main effects of Language ($\beta = 0.028$, $SE = 0.007$, $t = 3.94$, $p < .001$) and Trial type ($\beta = 0.191$, $SE = 0.014$, $t = 13.25$, $p < .001$) were both significant. Naming latency was shorter in L1 than in L2 ($M = 720 \pm 187$ ms and $M = 737 \pm 166$ ms, respectively) and shorter in single trials than in repeat trials ($M = 660 \pm 129$ ms and $M = 803 \pm 191$ ms,

respectively). There was a significant interaction between Language and Trial type ($\beta = 0.087$, $SE = 0.013$, $t = 6.89$, $p < .001$), suggesting that the mixing cost was larger in L1 ($M = 177$ ms; $z = 14.60$, $p < .001$) than in L2 ($M = 116$ ms; $z = 9.57$, $p < .001$). Although there was no significant three-way interaction among RSI, Language, and Trial type ($\beta = 0.005$, $SE = 0.025$, $t = 0.19$, $p = .850$), we further analyse to understand the relative magnitudes of mixing costs between L1 and L2. The interaction between Language and Trial type was significant in short RSI ($z = 5.01$, $p < .001$), indicating that the mixing cost was greater in L1 than in L2 ($M = 160$ ms and $M = 101$ ms, respectively). Similarly, in long RSI, the L1 mixing cost was significantly greater than the L2 mixing cost ($M = 195$ ms and $M = 132$ ms, respectively; $z = 4.73$, $p < .001$). These data indicate asymmetrical mixing costs in both the short and long RSI conditions (see Figure 2).

The accuracy analysis showed that the main effect of Trial type was marginally significant ($\beta = 6.436$, $SE = 3.325$, $z = 1.94$, $p = .053$), with higher accuracy in single trials than in repeat trials ($M = 99.9\% \pm 0.3\%$ and $M = 98.5\% \pm 1.7\%$, respectively). Apart from this, other effects were not significant ($ps > .1$).

2.5. Discussion

The aim of Experiment 1 was to examine the effect of RSI lengths on reactive language control and proactive language control across L1 (Chinese) and L2 (English) in spoken production. Results revealed significant switch costs in both languages, with the asymmetry in switch costs affected by RSI lengths. Specifically, there was an L1-L2 asymmetry in switch costs (which was greater in L1 than in L2) in the short RSI condition but a symmetry in switch costs (which was comparable between L1 and L2) in the long RSI condition. This finding is consistent with prior research demonstrating that unbalanced bilinguals inhibit the dominant L1 activation during L2 processing to a greater extent than they inhibit the L2 during L1 processing (Costa & Santesteban, 2004; Linck et al., 2012; Meuter & Allport, 1999; Philipp et al., 2007). Additionally, Experiment 1 found asymmetrical mixing costs in both short and long RSIs. There were robust mixing costs by comparing the repeat trials in the mixed-language context to the trials in the single-language context for each language, and the mixing cost was greater in L1 than in L2. This pattern aligns with previous studies (Christoffels et al., 2007; Prior & Gollan, 2011), suggesting that bilinguals recruit more proactive language control in L1 than in L2 during spoken language production (Ma et al., 2016).

3. Experiment 2

3.1. Participants

In this experiment, 56 Chinese–English bilinguals from South China Normal University participated (37 females; Mage = 21.00, SDage = 2.08). All participants[2] were non-English major students, right-handed, with normal or corrected-to-normal vision, and no history of neurological impairments or language disorders. They signed the written informed consent form. The local authority approved the experiment. Participants self-reported their proficiency in listening, speaking, reading, and writing on a 7-point scale for both Chinese and English and completed a lexTALE test. Participants were randomly assigned to the short and long RSI conditions for a cued language switching task in written naming. Again, treating the interaction effects as
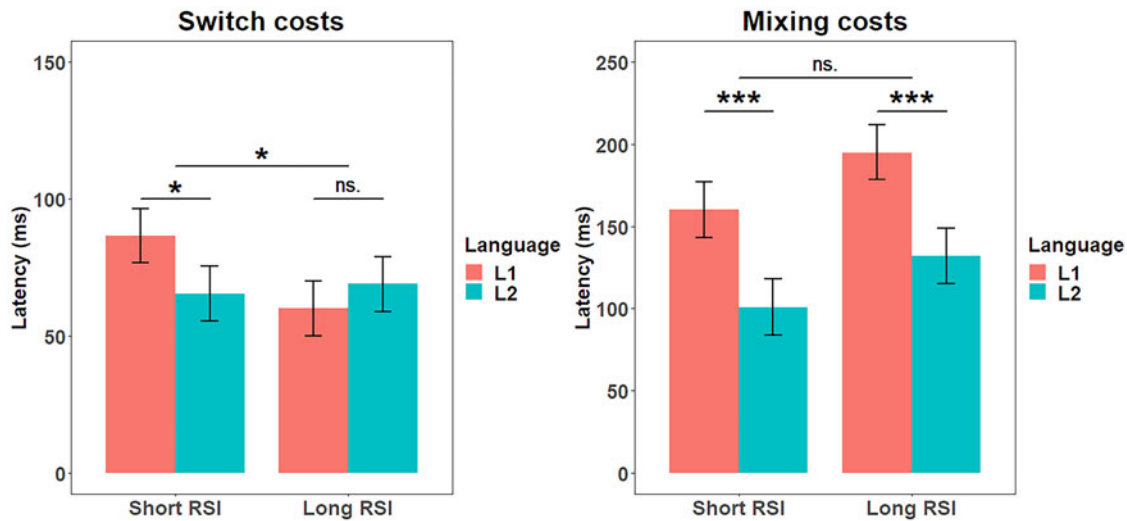
Figure 2. Mean switch costs and mixing costs for short RSI and long RSI conditions in Experiment 1. Error bars represent standard errors.

difference scores, we showed that, with $N = 28$ for each group, the experiment had a power of 83.6% to detect an expected effect size of $d_z = 0.8$.

As shown in Table 3, there were no significant differences in age, proficiency scores, L2 Age of Acquisition (AoA), and lexTALE test scores between the two groups (all $ps > .1$). Paired samples t-tests showed that proficiency scores for Chinese were significantly higher than for English in all four skills: listening ($t = 22.32$, $p < .001$), speaking ($t = 16.93$, $p < .001$), reading ($t = 14.75$, $p < .001$), and writing ($t = 9.96$, $p < .001$). These results indicate that the participants were unbalanced bilinguals with a dominant L1.

3.2. Materials

The materials were identical to those used in Experiment 1.

3.3. Task and procedure

The procedure was similar to that of Experiment 1, except that participants were instructed to write down the names corresponding to

the pictures on a WACOM Intuos A4 graphic tablet with a WACOM inking digitizer pen (Wacom, Japan). The experimental setup was the same as that in Experiment 1, with two single-language blocks and four mixed-language blocks. The interval from response to the next stimulus onset was 2000 ms for the short RSI condition and 3500 ms for the long RSI condition.

3.4. Results

We used the same data preprocessing and trimming procedures as in Experiment 1. For name latencies data, we first eliminated trials that were not recorded properly due to writing pen trigger failure (0.68% of all trials). Next, we removed these trials where a participant produced an incorrect word or used the wrong language, as well as the subsequent trial where the wrong language was used (3.78% of all trials). Then, we removed trials with latencies below 300 ms or above 2500 ms (1.14% of all trials). Lastly, we eliminated trials with latencies that were more than 2.5 SDs away from the participants' mean (2.84% of all trials). As in Experiment 1, we log-transformed latencies data to reduce the

Table 3. Participant characteristics in Experiment 2 (with SD in parentheses)

		Short RSI condition	Long RSI condition	*t*	*p*
	N (Female)	28 (18)	28 (19)		
	Age	21.36 (2.13)	20.64 (1.99)	1.29	.202
Chinese	Listening	6.50 (0.79)	6.43 (0.63)	0.37	.711
(L1)	Speaking	6.07 (1.02)	6.14 (0.89)	-0.28	.781
	Reading	6.14 (0.93)	6.11 (0.92)	0.15	.886
	Writing	5.64 (1.13)	5.43 (0.88)	0.79	.432
English	Listening	3.36 (0.91)	3.61 (0.69)	-1.16	.251
(L2)	Speaking	3.29 (1.18)	3.36 (0.78)	-0.27	.791
	Reading	4.39 (0.92)	4.14 (0.85)	1.06	.294
	Writing	3.50 (1.23)	3.79 (1.23)	-0.87	.389
	L2_AoA	7.75 (2.53)	8.04 (2.12)	-0.46	.649
	LexTALE	58.53 (7.96)	57.37 (7.24)	0.57	.571

skewness (pre-transformation skewness = 0.63 and post-transformation skewness = -0.33). For accuracy analyses, the error trials included using wrong words and language. In addition, as in Experiment 1, the data analysis used linear mixed-effects models and logistic mixed-effects models, and applied the same method and procedure to code variables and determine the optimal model. Table 4 presents the descriptive results.

Switch costs

The analysis of log latencies revealed the significant main effects of Language ($\beta = 0.029$, $SE = 0.012$, $t = 2.37$, $p = .034$) and Trial type ($\beta = 0.071$, $SE = 0.007$, $t = 10.68$, $p < .001$). Naming latency was shorter in L1 than in L2 ($M = 1046 \pm 261$ ms and $M = 1069 \pm 248$ ms, respectively), and shorter for repeat trials than for switch trials ($M = 1021 \pm 241$ ms and $M = 1095 \pm 264$ ms, respectively). Additionally, there was a significant interaction between RSI and Trial type ($\beta = 0.031$, $SE = 0.010$, $t = 3.11$, $p = .003$), indicating that the switch cost was larger in short RSI ($M = 95$ ms; $z = 10.37$, $p < .001$) than in long RSI ($M = 62$ ms; $z = 6.71$, $p < .001$). Moreover, the interaction between Language and Trial type was also significant ($\beta = 0.040$, $SE = 0.008$, $t = 4.98$, $p < .001$), implying that the switch cost was larger in L1 ($M = 98$ ms; $z = 11.76$, $p < .001$) than in L2 ($M = 58$ ms; $z = 6.60$, $p < .001$). The three-way interaction among RSI, Language, and Trial type was not significant ($\beta = 0.025$, $SE = 0.016$, $t = 1.58$, $p = .114$); however, as in Experiment 1, we still conducted a further analysis in order to shed further light on the relative magnitudes of switch costs between L1 and L2. The interaction between Language and Trial type was significant in short RSI ($z = 4.60$, $p < .001$), with a larger switch cost in L1 than in L2 ($M = 121$ ms and $M = 69$ ms, respectively). Similarly, in long RSI, the interaction between Language and Trial type reached significance ($z = 2.43$, $p = .015$), with a larger switch cost in L1 than in L2 ($M = 75$ ms and $M = 48$ ms, respectively). These findings suggested an asymmetry in the switch cost in both the short and long RSI conditions (see Figure 3).

The analysis of accuracy revealed a marginally significant main effect of RSI ($\beta = 0.434$, $SE = 0.239$, $z = 1.82$, $p = .069$), with higher accuracy in long RSI than in short RSI ($M = 97.8\% \pm 1.7\%$ and $M = 96.7\% \pm 2.4\%$, respectively). The significant main effect of Trial type ($\beta = 0.571$, $SE = 0.124$, $z = 4.60$, $p < .001$) suggested a higher accuracy in repeat trials than in switch trials ($M = 98.0\%$ $\pm 1.9\%$ and $M = 96.5\% \pm 2.9\%$, respectively). Moreover, there was a significant three-way interaction among RSI, Language, and Trial type ($\beta = 1.627$, $SE = 0.615$, $z = 2.64$, $p = .008$). Further analysis revealed that the interaction between Language and Trial type was significant in short RSI ($z = 2.42$, $p = .015$), with a larger switch cost in L1 than in L2 ($M = 3.4\%$ and $M = 0.4\%$, respectively); however, the interaction between Language and Trial type was not significant in long RSI ($z = -1.09$, $p = .275$), which indicated that the switch cost did not significantly differ between L1 and L2 ($M = 0.4\%$ and $M = 1.8\%$, respectively). These findings suggest the asymmetrical switch costs in the short RSI condition and the symmetrical switch costs in the long RSI condition.

Mixing costs

The analysis of log latencies revealed a marginally significant main effect of RSI ($\beta = 0.065$, $SE = 0.037$, $t = 1.72$, $p = .090$), with shorter latency in short RSI than in long RSI ($M = 927 \pm 226$ ms and $M = 983 \pm 226$ ms, respectively). The main effects of Language ($\beta = 0.051$, $SE = 0.015$, $t = 3.38$, $p = .003$) and Trial type ($\beta = 0.130$, $SE = 0.016$, $t = 8.21$, $p < .001$) were also significant. Naming latency was shorter in L1 than in L2 ($M = 934 \pm 227$ ms and $M = 977 \pm 226$ ms, respectively) and shorter for single trials than repeat trials ($M = 895 \pm 196$ ms and $M = 1021 \pm 241$ ms, respectively). The interaction among RSI, Language, and Trial type was not significant ($\beta = 0.005$, $SE = 0.027$, $t = 0.20$, $p = .840$). Further analysis was conducted to understand better the relative magnitudes of mixing costs between L1 and L2. There was no significant difference in mixing costs between L1 and L2 in either short RSI ($M = 128$ ms and $M = 132$ ms, respectively; $z = -0.33$, $p = .745$) or long RSI ($M = 129$ ms and $M = 126$ ms, respectively; $z = 0.07$, $p = .945$). These data suggested that the mixing costs were symmetrical in both RSI conditions (see Figure 3).

The analysis of accuracy found a marginally significant main effect of Trial type ($\beta = 6.525$, $SE = 3.859$, $z = 1.69$, $p = .091$), with higher accuracy in single trials than in repeat trials ($M = 99.7\% \pm 0.7\%$ and $M = 98.0\% \pm 1.9\%$, respectively). Apart from this, other effects were not significant ($ps > .1$). The detailed results of the accuracy analysis for Experiment 1 and Experiment 2 were reported in the Supplementary Material for interested readers.

3.5. Discussion

Experiment 2 examined the effect of RSI lengths on reactive and proactive language controls between L1 and L2 in written production. Firstly, in latency and accuracy data, it was observed that there were switch costs and mixing costs in written production. This empirically demonstrates that reactive language control, as measured by switch costs, and proactive language control, as measured by mixing costs, exist in written production, as in spoken production.

Secondly, latency data showed that there was an asymmetrical pattern of switch costs in both the short and long RSI conditions. These results were consistent with those of Ma et al. (2016), who observed asymmetrical switch costs at the RSIs of 500 ms, 800 ms, and 1500 ms in spoken production. Moreover, we found that accuracy manifested as asymmetrical switch cost in the short RSI condition and symmetrical switch cost in the long RSI condition; this modulation of RSI lengths on switch cost is in line with the finding of latency in Experiment 1.

Last but not least, this experiment revealed that the mixing cost was comparable between L1 and L2 in both short and long RSI conditions. This symmetry in the mixing costs in written

Table 4. Mean naming latency (SD) and accuracy (SD) of as a function of trial type across two response-stimulus intervals (RSIs) in Experiment 2

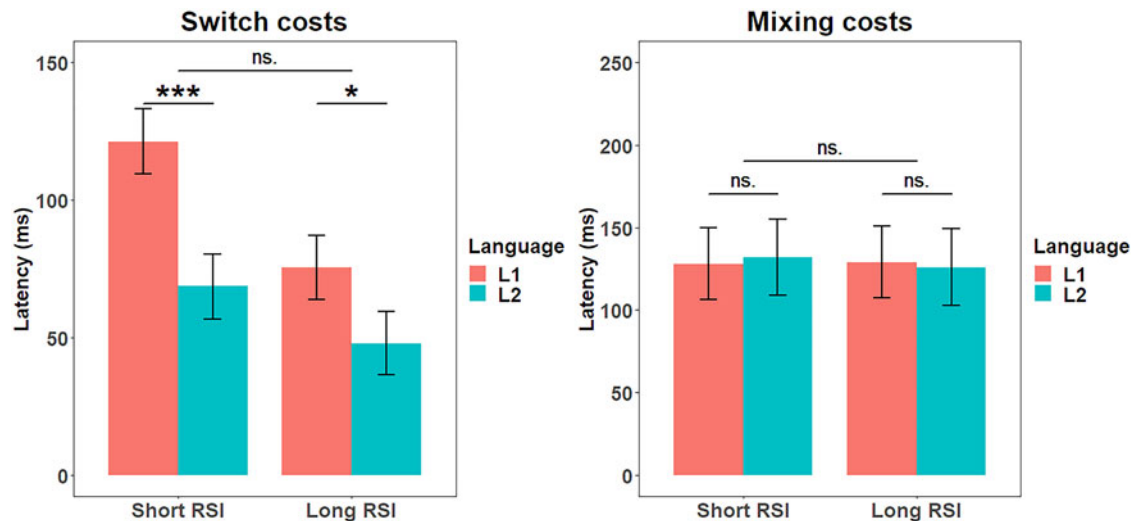|		Switch	Repeat	Single
Latency (ms)				
Short RSI	L1	1086 (277)	969 (233)	839 (195)
	L2	1089 (261)	1023 (240)	890 (190)
Long RSI	L1	1101 (265)	1029 (248)	908 (186)
	L2	1103 (250)	1061 (235)	942 (197)
Accuracy (%)				
Short RSI	L1	94.5 (4.3)	97.9 (2.3)	100.0 (0.0)
	L2	96.9 (4.3)	97.3 (3.3)	99.4 (1.7)
Long RSI	L1	97.7 (3.1)	98.1 (2.4)	99.9 (0.5)
	L2	96.9 (3.4)	98.7 (1.4)	99.6 (0.8)

Figure 3. Mean switch costs and mixing costs for short RSI and long RSI conditions in Experiment 2. Error bars represent standard errors.

production differs from the findings of spoken production in Experiment 1, which revealed an L1-L2 asymmetry in mixing costs in both RSIs. Therefore, in written production, bilinguals may engage modality-specific proactive control mechanisms to adjust the activation level of the two languages in single and mixed language contexts.

4. Additional analyses comparing the two modalities

The design and procedure of Experiment 1 and Experiment 2 were identical, with the exception of the production modality. To more clearly compare the relationship in bilingual control between spoken and written productions, we conducted between-experiment comparisons on log latencies and accuracy. We first conducted an independent sample t-test, which showed that there was no difference between the two groups of participants (i.e., response in spoken production and written production, respectively) in age, proficiency scores for Chinese and English in the four skills, L2 AoA, and lexTALE test scores ($ps > .1$).

In the main analyses reported below, all variables were coded using deviation contrasts, and the best-fitting model is reported. The models included Modality, RSI, Language, Trial type (switch costs analyses involved repeat and switch; mixing costs analyses involved single and repeat), and their interactions as fixed effects.

Switch costs
The results of the log latencies analysis showed a significant main effect of Modality ($t = 8.37$, $p < .001$) and a marginally significant main effect of RSI ($t = 1.82$, $p = .072$), showing shorter latency in spoken production than in written production and in short RSI than in long RSI. The main effect of Trial type reached significance as well ($t = 21.15$, $p < .001$), reflecting shorter latency in repeat trials than in switch trials. In addition, the two-way interaction between Modality and Language was significant as well ($t = 4.90$, $p < .001$), indicating that there was longer latency in L1 than in L2 in spoken production ($z = 1.90$, $p = .058$), but shorter latency in L1 than in L2 in written production ($z = -2.80$, $p = .005$). The significant interaction between RSI and Trial type ($t = 3.27$, $p < .001$) indicates a larger switch cost in short RSI than in long RSI. The significant interaction between Language

and Trial type ($t = 3.87$, $p < .001$) suggests that the switch cost being larger in L1 than in L2. Moreover, the significant three-way interaction among Modality, Language, and Trial type ($t = 2.64$, $p = .010$) indicates that there were symmetrical switch costs in spoken production but asymmetrical switch costs in written production overall. More importantly, the three-way interaction among RSI, Language, and Trial type was also significant ($t = 2.71$, $p = .008$). Further analysis found a significant interaction between Language and Trial type in short RSI ($z = 4.64$, $p < .001$) but not in long RSI ($z = 0.82$, $p = .412$), indicating asymmetrical switch costs in short RSI and symmetrical switch costs in long RSI. This was confirmed by analyses examining RSI modulation effects for spoken and written productions separately, showing asymmetrical switch costs in spoken (87 ms vs. 66 ms) and written productions (121 ms vs. 69 ms) when RSI was short.

The analysis of accuracy revealed a marginally significant main effect of RSI ($z = 1.87$, $p = .062$), with higher accuracy in long RSI than in short RSI. In addition, the main effects of Language ($z = 2.59$, $p = .010$) and Trial type ($z = 7.83$, $p < .001$) were also significant, with higher accuracy in L2 compared to L1 and in repeat trials than in switch trials. The interaction between Modality and Trial type reached marginal significance ($z = 1.73$, $p = .085$), with greater switch cost in spoken production than in written production. More importantly, the three-way interaction among RSI, Language, and Trial type was also significant ($z = 2.19$, $p = .029$), suggesting that there was a marginally significant interaction between RSI and Trial type in LI ($z = 1.89$, $p = 0.059$) but not in L2 ($z = -1.28$, $p = 0.201$). In addition, the four-way interaction among Modality, RSI, Language, and Trial type was marginally significant ($z = 1.88$, $p = .061$), meaning that there was a significant three-way interaction among RSI, Language, and Trial type in written production ($z = 3.06$, $p = .002$) but not in spoken production ($z = 0.21$, $p = .834$). Analyzing this three-way interaction in written production revealed that there were asymmetrical switch costs in short RSI (3.4% and 0.4%) and symmetrical switch costs in long RSI (0.4% and 1.8%) in writing. No other main effects or interactions were significant ($ps > .1$).

Mixing costs
The results of log latencies revealed the significant main effects of Modality ($t = 11.43$, $p < .001$) and RSI ($t = 2.38$, $p = .019$), with

shorter latency in spoken production compared to written production and in short RSI than in long RSI. Furthermore, the main effects of Language ($t = 6.02$, $p < .001$) and Trial type ($t = 14.02$, $p < .001$) were significant as well, showing shorter latency in L1 than in L2 and in single trials than in repeat trials. In addition, the marginally significant interaction between Modality and Language ($t = 1.75$, $p = .083$) showed that relative to spoken production ($z = 3.02$, $p = .003$), there was a greater magnitude of naming speed difference between L1 and L2 in written production ($z = 5.49$, $p < .001$). The significant interaction between Modality and Trial type ($t = 3.19$, $p = .002$) indicates greater mixing costs in spoken production ($z = 12.77$, $p < .001$) than in written production ($z = 8.66$, $p < .001$). The significant interaction between Language and Trial type ($t = 3.34$, $p = .005$) indicated that overall there were greater mixing costs in L1 ($z = 13.61$, $p < .001$) than in L2 ($z = 10.38$, $p < .001$). More importantly, the three-way interaction among Modality, Language, and Trial type was also significant ($t = 4.53$, $p < .001$), indicating that the interaction between Language and Trial type was significant in spoken production ($z = 5.29$, $p < .001$) but not in written production ($z = 0.28$, $p = .782$). That is, there were asymmetrical mixing costs in spoken production and symmetrical mixing costs in written production.

The analysis of accuracy revealed that there were no effects that reached significance ($ps > .1$)

5. General discussion

The current study aimed to investigate the universal and specific bilingual control mechanisms between spoken and written productions. To this end, unbalanced Chinese–English bilinguals performed a cued language switching task in spoken picture naming (Experiment 1) or written picture naming (Experiment 2), respectively. The results indicated that the reactive language control, as measured by switch costs, operates in a similar manner between speaking (Experiment 1) and writing (Experiment 2). However, the proactive language control, as measured by mixing costs, was found to operate in modality-specific manners. The implications of these findings are discussed in the following sections.

Similarities in bilingual language control between speaking and writing

The two experiments in the present study replicated some common observations regarding switch costs and mixing costs in both spoken and written productions. Specifically, the switch trials took longer naming latency and had lower accuracy than the repeat trials in both modalities, in line with previous research (Wong & Maurer, 2021). Additionally, the repeat trials in a mixed naming context had longer naming latency and lower accuracy than the trials in a single naming context in spoken production (Declerck et al., 2012; Prior & Gollan, 2011) and written production. Notably, to our best knowledge, this study is the first to show a mixing cost in written production, suggesting that sustained, proactive language control is at play in written production, adjusting the overall activation levels of the two languages (Gade et al., 2021).

Our analyses of naming latencies also showed an asymmetric pattern of switch costs in the short RSI condition for both spoken and written productions, with larger L1 switch costs than L2 switch costs. This finding is consistent with the assumption of the inhibitory control model (Green, 1998), according to which

the switch cost asymmetry arises because unbalanced bilinguals engage more effort in inhibiting their dominant L1 during L2 processing than in inhibiting L2 during L1 processing (Christoffels et al., 2007; Costa & Santesteban, 2004; Verhoef et al., 2009). Furthermore, previous research has classified the inhibition reflected in switch costs as reactive language control and the inhibition reflected in mixing costs as proactive language control (e.g., Declerck, 2020; Ma et al., 2016; Peeters & Dijkstra, 2018). Our findings suggest that this reactive language control applies to spoken and written productions, and that switching from L2 to L1 requires additional reactive language control to resolve residual inhibition of L1 (Ma et al., 2016; Mosca & de Bot, 2017).

In addition to the inhibitory control model, there are alternative accounts that attempt to explain the asymmetrical switch costs in unbalanced bilinguals (Philipp et al., 2007; Verhoef et al., 2009). Philipp et al. (2007) proposed the persistent activation account, arguing that this asymmetry in switch costs is due to top-down activation rather than inhibition. According to them, the activation of the naming language in the current trial carries over to the subsequent trial: since the non-dominant L2 is more strongly activated than the dominant L1, and this strong residual activation of the non-dominant L2 presents a barrier to the activation of the dominant L1, resulting in a greater switch cost to the dominant language. Our study showed that, as RSI lengthened, the switch cost asymmetry was reduced or even disappeared. This result can be, in light of the persistent activation account, attributed to the declining L2 residual activation as RSI increases.

The L1 repeat benefit account was proposed by Verhoef et al. (2009). They observed the asymmetrical switch costs in short CSI condition and the symmetrical switch costs in long CSI condition, and found that the preparation interval effects (i.e., the faster naming speed in long CSI than in short CSI) were present in all conditions (i.e., L1 switch, L2 switch, and L2 repeat trials) except for the L1 repeat trials. Thus, they argue that cross-language interference exists in all conditions except for the L1 repeat trials, resulting in asymmetric switch costs. For the spoken production in the present study, we found another interval effect in the mixed-language context–that is, with shorter latency for short RSI compared to long RSI–and this interval effect was observed in all conditions except for the L1 switch trials. Following a similar logic as in Verhoef et al. (2009), we can come to a different conclusion that language competition is absent in the special case of the L1 switch trials. Therefore, our findings didn't support this L1 repeat benefit account. In fact, the asymmetrical switch costs may be jointly determined by the four trial types (i.e., L1 switch, L1 repeat, L2 switch, and L2 repeat), and the RSI modulation effect on switch cost asymmetry in spoken production may be a result of the dissipation time of the task-set/language-set competition. The long waiting time for passive dissipation of the language-set competition after the response helps reduce the need for stronger reactive language control of L1.

In addition to the findings discussed previously, the present study provides evidence that reactive language control was similarly modulated by RSI between spoken and written productions. Results from both latencies in spoken production and from accuracy in written production seem to indicate a similar pattern. In terms of latencies, there were asymmetric switch costs with short RSI and symmetric switch costs with long RSI in spoken production. Notably, in terms of accuracy, RSI lengths modulated the switch cost asymmetry in written production, with

asymmetric switch costs when RSI was short but symmetric switch costs when RSI was long. These findings suggest a similar modulation of RSI on reactive language control in both speaking and writing. This result extends the finding of Ma et al. (2016) that switch cost asymmetry was less prominent with longer RSIs (800 and 1500 ms) than with shorter RSI (500 ms). The current study used increased RSIs (short = 2000 ms vs. long = 3500 ms) and showed that, as the waiting time after a response increases, the switch costs show a symmetrical pattern, indicating that with the increase of RSI, the difference in reactive language control between L1 and L2 decreases or even disappears.

Furthermore, it is worth noting that our comparative analyses of latencies revealed symmetrical switch costs in spoken production but asymmetrical switch costs in written production. Previously, Wong and Maurer (2021) found similar switch costs for speaking and writing. Though the study did not directly test for the (a)symmetry of switch costs, it can be inferred from the results that both speaking and writing exhibited symmetrical switch costs. In addition, a recent study by Roembke et al. (2023) also revealed symmetrical switch costs in typewritten production across four experiments. In fact, a recent meta-analysis by Gade et al. (2021) failed to find substantial evidence for an asymmetry in the switch cost, with a similar number of studies observing either asymmetrical switch costs (e.g., Peeters & Dijkstra, 2018; Philipp et al., 2007) or symmetrical switch costs (e.g., Ivanova & Hernandez, 2021; Mosca & Clahsen, 2016). Therefore, considering these findings, it is plausible that the (a) symmetry of switch cost in written production could be task-dependent or exhibit significant variability. The results of this study indicated that the asymmetry in switch costs varied according to the length of RSI and demonstrated symmetrical switch costs in spoken production overall. However, in written production, this asymmetry in latencies was unaffected by RSI lengths and displayed asymmetrical switch costs in short and long RSIs. This distinction can be attributed to the fact that writing demands more intricate coordination of visual perceptual and motor control processes compared to speaking, resulting in potentially longer production times (Breining & Rapp, 2019; Perret et al., 2014). Also for this reason, the RSI of the written production was set to be 1-2 seconds longer than that of the spoken modality in previous monolingual studies (e.g., Damian et al., 2011; Qu et al., 2021). Collectively, we speculate that it may be that the range of RSI lengths utilized in this study may not have been sufficiently long to detect the modulation of switch cost asymmetry over time in written production.

Differences in bilingual language control between speaking and writing

In the present study, unbalanced bilinguals exhibited different patterns of mixing costs in spoken production and written production. Language mixing costs serve as an indicator of proactive language control and reflect global inhibitory processing for cross-language interference (Declerck, 2020; Ma et al., 2016).

Experiment 1 revealed that mixing costs were greater in L1 than in L2 in both the short and the long RSI conditions in spoken production (i.e., asymmetrical mixing costs). This asymmetrical pattern in mixing costs is consistent with prior studies (Christoffels et al., 2007; Jylkkä et al., 2018; Prior & Gollan, 2011; Timmer et al., 2017), suggesting that unbalanced bilinguals need to recruit more proactive language control in L1 compared to L2 in spoken production. Additionally, Experiment 1 did not observe a modulation by RSI on the mixing cost asymmetry. In contrast, Experiment 2 showed that, for writing, the mixing costs did not differ between L1 and L2 in both the short and the long RSI conditions (i.e., symmetrical mixing costs). This symmetrical pattern of mixing costs stands in contrast with the pattern observed in Experiment 1 and some prior studies (Prior & Gollan, 2011; Timmer et al., 2017). Additionally, the size of the mixing cost was not modulated by RSI. Our combined analysis further confirmed that speaking and writing exhibit different patterns of mixing costs, with unbalanced bilinguals showing greater proactive language control in L1 during speaking and similar levels of proactive language control in L1 and L2 during writing.

The way in which language mixing affects performance may have implications for proactive control in spoken and written productions. Specifically, in speaking, naming latency was shorter in L1 than in L2 in the single language context for single trials, indicating a higher baseline activation level for L1 than for L2. However, for repeat trials in the mixed language context, there was no difference in naming latency between L1 and L2 in speaking, suggesting that the L1 naming advantage is reduced or even disappears in the mixed language context. This supports the theory that L1 is globally inhibited in the mixed context to facilitate L2 spoken production (Christoffels et al., 2007). In writing, L1 naming was faster than L2 naming for single and repeat trials, which suggests that unbalanced bilinguals can maintain the speed advantage of L1 handwriting in single and mixed language contexts. Additionally, the combined analysis suggests that response latency may have inconsistent effects on the size of the switch costs and mixing costs. In both single and mixed language contexts, there was a longer latency in writing than in speaking. Notably, the size of switch costs was comparable between the two production modalities but a smaller mixing cost in writing than in speaking; this supports the idea to a certain extent that the switch costs are similar between modalities, but the mixing costs differ.

Finally, there were several potential limitations to the present study. First of all, the number of participants in the current study is relatively small, which might have compromised the statistical power of our experiments, especially in the between-experiment comparison. Indeed, post-hoc simulation-based power analyses (Kumle et al., 2021) showed that, for switch costs, the statistical power for the three-way interaction on log latencies analysis in Experiment 1 was only 68.9%, and the statistical power for the three-way interaction on accuracy analysis in Experiment 2 was 83.4%. Hence, one should exercise caution when interpreting the three-way interaction and, to some extent, the two-way interactions as well. However, the analyses indicated that there was sufficient statistical power (98.0%) to detect the three-way interaction among Modality, Language, and Trial for mixing costs. Detailed statistical power can be seen in Supplementary Material. Future research can directly compare language control between speaking and writing using a within-subject design.

Another potential limitation relates to the benefit of repeating the same cue, which may involve some additional processing that has nothing to do with language control. In the present study, the cued language switching task used 1:1 cue-to-language mappings (i.e., one cue represents each language); language repetitions are triggered by presenting the same cues consecutively, and language switches by presenting the different cues in succession. It has been suggested that cue switches could elicit additional cue-switch costs (Forstmann et al., 2007; Heikoop et al., 2016). Therefore,

future studies can consider using 2:1 cue-to-language mapping (i.e., two cues represent each language) or other ways to further study the pure language control effect.

In conclusion, the present study revealed modality-general and modality-specific bilingual control mechanisms between speaking and writing in unbalanced bilinguals. In particular, our findings revealed that greater reactive language control was observed in the dominant L1 compared to the less dominant L2 in both modalities when the RSI was short. In addition, we found stronger proactive language control in L1 than in L2 in speaking, but no difference in proactive language control between the two languages in writing. These findings indicated that the reactive language control is modality-general between speaking and writing, while the proactive language control is modality-specific.

Notes

[1] The sample size of participants was determined based on the (post hoc) power analysis. Four out of the twenty-eight participants in each condition were tested at a later point in time in Experiment 1.

[2] The sample size of participants was determined based on the (post hoc) power analysis. Four out of the twenty-eight participants in each condition were tested at a later point in time in Experiment 2.

References

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neuro-cognition of language representation and control. *Journal of Neurolinguistics*, 20(3), 242–275.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.

Bialystok, E., Craik, F. I., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(6), 1341–1354.

Bobb, S. C., & Wodniecka, Z. (2013). Language switching in picture naming: What asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning. *Journal of Cognitive Psychology*, 25(5), 568–585.

Breining, B., & Rapp, B. (2019). Investigating the mechanisms of written word production: insights from the written blocked cyclic naming paradigm. *Reading and Writing*, 32, 65–94.

Broos, W. P., Bencivenni, A., Duyck, W., & Hartsuiker, R. J. (2021). Delayed picture naming in the first and second language. *Bilingualism: Language and Cognition*, 24(2), 389–400.

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, 1147, 192–208.

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(4), 491–511.

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(5), 1057–1074.

Damian, M. F., Dorjee, D., & Stadthagen-Gonzalez, H. (2011). Long-term repetition priming in spoken and written word production: evidence for a contribution of phonology to handwriting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(4), 813–826.

Declerck, M. (2020). What about proactive language control?. *Psychonomic Bulletin & Review*, 27(1), 24–35.

Declerck, M., & Koch, I. (2023). The concept of inhibition in bilingual control. *Psychological Review*, 130(4), 953.

Declerck, M., & Philipp, A. M. (2015). A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review*, 22, 1630–1645.

Declerck, M., Koch, I., & Philipp, A. M. (2012). Digits vs. pictures: The influence of stimulus type on language switching. *Bilingualism: Language and Cognition*, 15(4), 896–904.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.

Fink, A., & Goldrick, M. (2015). Pervasive benefits of preparation in language switching. *Psychonomic Bulletin & Review*, 22, 808–814.

Forstmann, B. U., Brass, M., & Koch, I. (2007). Methodological and empirical issues when dissociating cue-related from task-related processes in the explicit task-cuing procedure. *Psychological Research*, 71, 393–400.

Gade, M., Declerck, M., Philipp, A. M., Rey-Mermet, A., & Koch, I. (2021). Assessing the Evidence for Asymmetrical Switch Costs and Reversed Language Dominance Effects–A Meta-Analysis. *Journal of Cognition*, 4(1), 55.

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.

Gould, J. D., & Boies, S. J. (1978). Writing, dictating, and speaking letters. *Science*, 201(4361), 1145–1147.

Grange, J. A., & Cross, E. (2015). Can time-based decay explain temporal distinctiveness effects in task switching?. *The Quarterly Journal of Experimental Psychology*, 68(1), 19–45.

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

Hanulová, J., Davidson, D. J., & Indefrey, P. (2011). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Language and Cognitive Processes*, 26(7), 902–934.

Heikoop, K. W., Declerck, M., Los, S. A., & Koch, I. (2016). Dissociating language-switch costs from cue-switch costs in bilingual language switching. *Bilingualism: Language and Cognition*, 19(5), 921–927.

Horoufchin, H., Philipp, A. M., & Koch, I. (2011). The dissipating task-repetition benefit in cued task switching: Task-set decay or temporal distinctiveness?. *Journal of Experimental Psychology: Human Perception and Performance*, 37(2), 455–472.

Ivanova, I., & Hernandez, D. C. (2021). Within-language lexical interference can be resolved in a similar way to between-language interference. *Cognition*, 214, 104760.

Jylkkä, J., Lehtonen, M., Lindholm, F., Kuusakoski, A., & Laine, M. (2018). The relationship between general executive functions and bilingual switching and monitoring in language production. *Bilingualism: Language and Cognition*, 21(3), 505–522.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching – A review. *Psychological Bulletin*, 136(5), 849–874.

Koch, I., & Allport, A. (2006). Cue-based preparation and stimulus-based priming of tasks in task switching. *Memory & Cognition*, 34(2), 433–444.

Kumle, L., Võ, M. L. H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research Methods*, 53(6), 2528–2543.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13). https://doi.org/10.18637/jss.v082.i13

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.

Lenth, R. V. (2022). Emmeans: Estimated Marginal Means, aka Least-Squares Means, https://CRAN.R-project.org/package=emmeans.

Linck, J. A., Schwieter, J. W., & Sunderman, G. (2012). Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism: Language and Cognition*, 15(3), 651–662.

Liu, C., Jiao, L., Wang, Z., Wang, M., Wang, R., & Wu, Y. J. (2019). Symmetries of bilingual language switch costs in conflicting versus

non-conflicting contexts. *Bilingualism: Language and Cognition*, 22(3), 624–636.

Ma, F., Li, S., & Guo, T. (2016). Reactive and proactive control in bilingual word production: An investigation of influential factors. *Journal of Memory and Language*, 86, 35–59.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, 40(1), 25–40.

Mosca, M., & Clahsen, H. (2016). Examining language switching in bilinguals: The role of preparation time. *Bilingualism: Language and Cognition*, 19(2), 415–424.

Mosca, M., & de Bot, K. (2017). Bilingual language switching: Production vs. recognition. *Frontiers in psychology*, 8, 934.

Muylle, M., Van Assche, E., & Hartsuiker, R. J. (2022). Comparing the cognate effect in spoken and written second language word production. *Bilingualism: Language and Cognition*, 25(1), 93–107.

Peeters, D., & Dijkstra, T. (2018). Sustained inhibition of the native language in bilingual language production: A virtual reality approach. *Bilingualism: Language and Cognition*, 21(5), 1035–1061.

Perret, C., Bonin, P., & Laganaro, M. (2014). Exploring the multiple-level hypothesis of AoA effects in spoken and written object naming using a topographic ERP analysis. *Brain and Language*, 135, 20–31.

Philipp, A. M., Gade, M., & Koch, I. (2007). Inhibitory processes in language switching: Evidence from switching language-defined response sets. *European Journal of Cognitive Psychology*, 19(3), 395–416.

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.

Qu, Q., Feng, C., & Damian, M. F. (2021). Interference effects of phonological similarity in word production arise from competitive incremental learning. *Cognition*, 212, 104738.

Roembke, T. C., Koch, I., & Philipp, A. M. (2023). Language Switching when Writing: The Role of Phonological and Orthographic Overlap. *Quarterly Journal of Experimental Psychology*, Advance online publication. https://doi.org/10.1177/17470218231183706

Tainturier, M. J. (2019). A theory of bilingual spelling in alphabetic systems. *In Spelling and Writing Words: Theoretical and methodological advances*. Brill.

Timmer, K., Grundy, J. G., & Bialystok, E. (2017). Earlier and more distributed neural networks for bilinguals than monolinguals during switching. *Neuropsychologia*, 106, 245–260.

Verhoef, K., Roelofs, A., & Chwilla, D. J. (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. *Cognition*, 110(1), 84–99.

Wong, W. L., & Maurer, U. (2021). The effects of input and output modalities on language switching between Chinese and English. *Bilingualism: Language and Cognition*, 24(4), 719–729.

Zhang, Q., & Yang, Y. (2003). The determiners of picture-naming latency. *Acta Psychologica Sinica*, 35(4), 447–454. (In Chinese)