OPEN ACCESS

Australian Journal of Applied Linguistics

ISSN 2209-0959

https://www.castledown.com/journals/ajal/

Australian Journal of Applied Linguistics, 9, 102651 (2026)
https://doi.org/10.29140/ajal.2026.102651

The role of inhibitory control in the phonological processing of bilinguals during language switching

BAO SHU YI

Nanjing University of Posts and Telecommunications, China

Abstract

Bilingual individuals demonstrate the capacity to choose the intended language while inhibiting interference from unintended languages, sparking debates on the cognitive mechanism respon-sible for this successful selection. While most studies have focused on inhibition at the lemma level, fewer have delved into the cognitive mechanisms of language switching at the phonological level, with even fewer exploring the direct relationship between inhibition ability and switch costs at the phonetic level. This study aimed to investigate the occurrence of inhibitory control at the phonetic level during speech production when transitioning between Chinese and English. Employing a mixed design with both within-subjects and between-subjects components, the study involved switch and non-switch trials in both L1-L2 and L2-L1 transitions, across three groups categorized by inhibition ability levels (high, mid, and low). The investigation unfolded in two main parts: firstly, scrutinizing the impact of language switching on specific phonetic features in a cued picture naming task; and secondly, exploring the potential direct link between inhibition ability and switch costs at the phonetic level, specifically focusing on the voice onset time of stops. The results unveiled asymmetrical switch costs in the voice onset time of stops when transitioning into the dominant language (L1) for Chinese-English bilinguals, indicating a robust manifestation of inhibitory control at a specific linguistic juncture. Furthermore, a notable correlation emerged between inhibition abilities and switch costs (VOT cost), showcasing that stronger inhibition abilities correlated with reduced costs in voice onset time. Thus, the study suggests that inhibitory control may be imperative in bilingual language tasks, aiding in the suppression of cross-linguistic phonetic influences, and that enhanced inhibition abilities could mitigate the impact of the other language during language switching for Chinese and English bilinguals.

Keywords: Bilinguals, language switch, inhibitory control, switch cost, VOT of stops

Introduction

In the realm of second language (L2) processing and acquisition, it is widely recognized that adult learners exhibit considerable variability in their outcomes (Liu, 2024). Traditionally, explanations for this variability have focused on language proficiency or other linguistic aspects (Topal, 2024). Various internal psychological factors among learners have been put forth to explain such differences, including cognitive factors like working memory (Skehan, 2009; Schwieter & Wen, 2022) and inhibition (Lev-Ari & Peperkamp, 2014; Darcy *et al*., 2016), as well as affective factors such as motivation (Moyer, 1999) and personality traits (Kerz & Wiechmann, 2017). Recent research has highlighted significant correlations between language processing and cognitive factors (Kidd et al., 2018). Among these cognitive factors, working memory has received the most attention in the context of L2 processing (Wen et al., 2019; Dong & Li, 2020). However, another cognitive ability, inhibitory control, a mental control of the bilingual lexico-semantic systems in which bilinguals are able to select and activate the intended language and suppress the unintended language according to the specific situations, plays a crucial role as a cognitive process and a facet of executive function that enables individuals to control their impulses and override their natural responses to stimuli (Bajo et al., 2021; Li et al., 2021). Despite its importance, the role of inhibitory control in language processing, especially in bilingual language processing, remains understudied.

Green (1998) was the first to propose the Inhibitory Control Model (ICM), linking inhibitory control to language processing. According to the ICM, inhibitory control serves as a mental mechanism through which bilinguals can choose and activate their intended language while suppressing their unintended language based on the specific context. This model predicts asymmetrical switch costs (the measurable decline in processing speed or accuracy observed when bilinguals alternate between languages, compared to trials where they continue using the same language, reflecting the cognitive effort required to inhibit the previously active language system), with larger costs observed when switching to the more dominant language (L1).

The assumption regarding asymmetric switch costs was first tested through the language switching paradigm introduced by Meuter and Allport (1999). Their study revealed that participants exhibited switch costs, with slower naming latencies on switch trials compared to non-switch trials. Additionally, the results supported the predicted asymmetry, showing larger switch costs when switching to the dominant L1 compared to the non-dominant L2. Since Meuter and Allport's influential research, investigating the presence of asymmetric switch costs in the cued language switching paradigm has become a fundamental test for exploring the role of inhibitory control in bilingual language production (Linck et al., 2011; Schwieter & Sunderman, 2008; Verhoef, Roelofs & Chwilla, 2009; Yang, 2013). Although not all the studies mentioned above offer conclusive evidence for inhibition, there is growing support indicating the involvement of inhibitory processes during language switching. It is widely acknowledged that bilingual language processing likely relies, at least to some extent, on inhibitory control (Declerck & Koch, 2023). However, questions still persist regarding what exactly is being inhibited in bilingual language processing and the specific locus of this inhibition. Given that language processing, encompassing production and comprehension, involves distinct processing stages, each stage may play a critical role in language control. The Inhibitory Control Model (Green, 1998) initially proposed a role of inhibition at the lemma level (abstract, non-phonological representation). The potential role of inhibition at the phonological level during language processing has also attracted significant attention from researchers.

Research investigating the role of inhibitory control at the phonological level in bilingual language processing has gained considerable attention in recent years. A number of studies have delved into this area, including those conducted by Antoniou et al. (2012), Declerck et al. (2015), Filippi et al. (2014), Levy et al. (2007), Lev-Ari & Peperkamp (2013), Olson (2013), and Darcy et al. (2016). Early evidence supporting the involvement of inhibitory control at the phonetic level in bilingual language processing emerged from studies utilizing a bilingual retrieval practice paradigm, such as those by Levy et al. (2007) and Yang et al. (2016). They both reported the presence of asymmetrical switch costs at the phonetic level in a cross-language retrieval forgetting task. Subsequent studies that manipulated phonological characteristics in language switching tasks, particularly those focusing on cognate status, have further contributed to our understanding (Declerck et al., 2012; Filippi et al., 2014; Thomas & Allport, 2000). While investigations into the impact of cognates on switch costs have yielded noteworthy findings, there has been some variation in the observed switch cost patterns between studies using cognates and non-cognates. Studies implementing cognates or non-cognates in distinct blocks have reported smaller switch costs with cognates, whereas those employing cognates within the same block have shown larger switch costs for cognates compared to non-cognates (Declerck et al., 2015). Broersma (2011) proposed an explanation for the smaller switch costs with cognates when implemented in separate blocks, suggesting a persisting effect on the phonological level where switch trial interference was reduced after cognates compared to non-cognates.

The notion of persisting effects on the phonological level has been supported by the work of Declerck & Philipp (2015), who investigated the impact of phonological overlap across trials. Their findings suggest that phonological manipulation can influence persisting effects in language switching, providing further evidence for the involvement of inhibitory control in phonological processing during bilingual language tasks.

The recent focus on language switching effects on accents in bilingual language processing has shed light on how language switches impact phonological processing. Studies, such as those by Goldrick et al. (2014) and Olson (2013), have reported intriguing findings regarding accent changes in switch trials compared to repetition trials among Spanish-English bilinguals. These changes in accent highlight the influence of language switching on phonological processing. However, not all studies have consistently supported the idea of inhibitory control operating at the phonological level. Linck (2008), for example, did not find inhibitory effects in his study, indicating that not all tasks may engage phonological processing to the same extent. Similarly, Declerck et al. (2013) did not observe switch cost differences based on the phonetic characteristics of words, suggesting that inhibition may not always manifest at the phonetic level as previously theorized. Furthermore, the existing research predominantly focuses on words rather than specific phonetic properties in bilingual speech production. As a result, there is a gap in understanding how language switching may influence specific phonological processes, such as producing different types of stops (e.g., bilabial, dental, velar), in bilingual contexts.

To address this gap, the present study aims to investigate whether inhibitory control operates at specific phonetic levels, such as stops, when switching between Chinese and English. By examining the effects of language switching on these specific phonological processes, the study intends to provide insight into the nature of inhibitory control in bilingual phonological processing.

The Present Study

The two parts of the study to investigate inhibitory control at the phonetic level in speech production when switching between two languages are crucial for gaining insights into bilingual language processing. In the first part of the study, using a cued picture naming task to examine the effect of language switching on specific phonetic features provides a direct way to assess how switches between languages influence phonological processing. By focusing on phonetic properties such as stops (e.g., bilabial, dental, velar), the study can reveal how these specific aspects of speech production are affected by language switches. This approach allows for a detailed analysis of the phonetic level of inhibitory control in bilingual language tasks. In the second part of the study, exploring the relationship between inhibition ability and switch cost or mixed cost at the phonetic level can provide valuable insights into the underlying mechanisms of language switching effects on phonological processing. By investigating whether individual differences in inhibitory control correlate with performance measures related to phonetic processing during language switches, the study can elucidate how inhibitory processes operate at the phonetic level in bilingual speech production.

The investigation into the language switching effect on specific phonetic features, particularly focusing on voice onset time (VOT) of stops in a cued picture-naming task, adds a nuanced dimension to understanding the impact of language switches on phonetic production. In this study, analyzing VOT of stops is a precise and effective way to assess phonetic differences between switched and non-switched tokens. VOT, as defined by Lisker and Abramson (1964), serves as a critical measure to distinguish voicing and aspiration of stops in speech production. The temporal interval from the release of an initial stop to the onset of glottal pulsing for the following vowel provides insights into the timing and coordination of speech sounds. By categorizing VOT values into voicing lead, voicing short lag, and voicing long lag based on the timing of voicing relative to the release of the stop, the study can capture subtle variations in phonetic production affected by language switching. Differences in VOT values between switches and non-switches tokens can indicate how language switches influence the phonetic realization of stops in speech production. By integrating VOT analysis into the cued picture-naming task, the study aims to uncover how language switching impacts phonetic features at a detailed level of speech production. This approach provides a quantitative and objective measure to examine the influence of language switches on phonetic processing, contributing to a deeper understanding of bilingual phonological mechanisms. Thereby, Keating (1984) proposed a three-categorization of the stops: voiced, voiceless unaspirated, and voiceless aspirated. In English and Chinese, based on the place of articulation, the stops can be divided into the same three groups; bilabial (/p/, /b/ in English, /p/, /p^h/ in Chinese), alveolar (/t/, /d/) in English, (/t/, /t^h /) in Chinese, velar (/k/, /g/ in English, /k/, /k^h/ in Chinese). However, according to manner of articulation, English and Chinese differ in the stop system in that voicing distinguishes English stops into voiced (/b/, /d/, /g/) and voiceless (/p/, /t/, /k/) stops while in Chinese, all the stops are voiceless and aspiration which distinguishes Chinese stops into unaspirated (/p/, /t/, /k/) and aspirated (/p ^{h}/, /t ^{h} /, /k ^{h}/) ones (Zheng & Li, 2007). Moreover, both Chinese and English voiceless aspirated stops are long lag, but in English, voiced fully voiced stops are voicing lead, their VOT is negative, and the VOT of voiceless unaspirated stops is zero or short lag, whereas in Chinese, only unaspirated stops belong to a short lag category.

The utilization of VOT in examining stop-voicing distinction as a measure of phonetic effects in language switching research is grounded in several compelling reasons. Firstly, VOT is a well-established and reliable phonetic cue for categorizing stop consonants across languages, making it a universal tool for studying phonetic distinctions (Cho & Ladefoged, 1999; Lisker & Abramson, 1964). Secondly, previous studies on the phonetic effects of code-switching have successfully used VOT to investigate phonetic realizations at language switches, highlighting its importance in capturing linguistic changes (Grosjean & Miller, 1994; Antoniou et al., 2011; Olson, 2013). Thirdly, the widespread usage of VOT as a measure of stop-voicing distinction in both Chinese and English in previous research facilitates cross-study comparisons and enhances the interpretability of findings (Dai, 2012; Fowler et al., 2008; Li & Dong, 2014; Lisker & Abramson, 1964; Ran & Shi, 2007; Wang, 2009). The cued picture naming paradigm employed in this experiment holds significance in the language switching literature due to its ability to isolate language switches, remove discursive constraints, and provide a clear view of phonetic transfer mechanisms (Declerck & Phillip, 2015). By presenting visual cues to prompt naming in specific languages, this paradigm enables researchers to study language switches in a controlled setting that minimizes external influences and facilitates a focused examination of phonetic processes. Although commonly used in lexical separation studies, its application at the phonetic level has been relatively limited (Olson, 2013). In this study, the blocked and mixed naming contexts, along with the inclusion of switch and non-switch trials in both languages, offer a comprehensive exploration of inhibitory control mechanisms during language switching. By analyzing VOT differences between switch and non-switch trials, mixed and blocked naming conditions, and L2-to-L1 and L1-to-L2 switch orders, the study aims to uncover asymmetrical VOT patterns indicative of inhibitory processes. By examining how VOT values are affected by language switches, the research aligns with the inhibitory control model's predictions. It seeks to shed light on the differential interference between languages during switching.

Exploring the relationship between language switch cost and inhibition abilities is also a significant aspect of the study, aiming to investigate whether a strong correlation exists between performance on the Stroop task, a classic measure of inhibition, and the phonetic effects observed during language switching. The Stroop task is a well-established paradigm for evaluating inhibitory control by requiring individuals to respond to one stimulus dimension while suppressing interference from another, typically involving conflicting information (Stroop, 1935; Simon & Rudell, 1967). In bilingual research, the Stroop task has been utilized to assess inhibitory abilities and has shown relevance in understanding language processing and cognitive control in bilingual individuals (Linck et al., 2011; Bialystok et al., 2005; Coderre, 2010; Marian et al., 2012). By incorporating the arrow-based spatial conflict task, a variant of the Stroop task into the present experiments and correlating participants' performance on this task with VOT mean differences in language-switch tokens, the study aims to shed light on the role of inhibition in phonetic processing during language switches. The hypothesis posits that individuals with stronger inhibitory control abilities may exhibit reduced phonetic switch costs in terms of VOT differences, indicating a more efficient separation between phonetic systems and smoother transitions between languages. Establishing a correlation between inhibitory abilities and VOT differences during language switching would provide compelling evidence for the involvement of inhibition mechanisms in regulating phonetic processes and facilitating language switches between distinct phonetic systems.

Method

Participants

In the study, 44 graduate students participated, with 39 of them self-reporting as right-handed and five as left-handed. All participants were native Chinese speakers with an average age of 24.11 years. They had never visited English-speaking countries and had learned English as a foreign language for an average of 12.13 years. On average, the participants demonstrated a low level of English proficiency, as indicated by a college English test for Chinese students, with a mean score of 2.73. Chinese was identified as their dominant language, and their proficiency in Chinese language skills (listening, speaking, and writing) was significantly higher than their English proficiency. Specifically, the differences in proficiency levels between Chinese and English were found to be statistically significant with t-values of 9.89, 7.21, 10.04, and 7.35 for listening, speaking, reading, and writing skills, respectively, all with p-values below 0.001.

This indicates a considerable proficiency gap between the participants' dominant language (Chinese) and their foreign language (English) across various language competencies. More information about the participants was presented in Table 1 as follows:

Table 1. *Descriptive statistics of the participants' information*

	N	Range	Minimum	Maximum	Mean	SD	Skewness		Kurtosis	
Age	44	5	22	27	24.11	1.104	.200	.357	.158	.702
Age of LE	44	13	7	20	12.14	2.530	1.230	.357	2.132	.702
CET	44	4	1	5	2.73	.997	−.151	.357	−.405	.702
C-listening	44	2	3	5	4.34	.645	−.457	.357	−.625	.702
C-reading	44	3	2	5	4.14	.765	−.566	.357	−.036	.702
C-speaking	44	2	3	5	4.02	.628	−.016	.357	−.312	.702
C-writing	44	3	2	5	3.82	.843	.119	.357	−1.100	.702
E-listening	44	4	1	5	2.48	1.000	.212	.357	−.346	.702
E-level	44	4	1	5	2.70	.765	−.085	.357	1.593	.702
E-reading	44	3	2	5	3.27	.758	.161	.357	−.182	.702
E-speaking	44	4	1	5	2.55	.901	.058	.357	.268	.702
E-writing	44	3	1	4	2.86	.765	−.739	.357	.815	.702

Age of LE = age of learning English; CET = College English testing; C-listening = Chinese listening, C-reading = Chinese reading; C-speaking = Chinese speaking; C-writing = Chinese writing; E-listening = English-listening; E-level = English level; E-reading = English reading; E-speaking = English speaking; E-writing = English writing.

Material, Procedure, and Design

Stimuli used in this experiment were selected from a self-built picture database containing 108 black and white pictures of the same size (300 mm × 300 mm). Sixty college students who did not participate in this experiment were recruited to test the predictability and familiarity of the names associated with these pictures. There are 90 pictures with consistent naming predictability (mean: 0.87 ± 0.12) and familiarity (4.09 ± 0.36) were selected for naming in the cued picture naming task to ensure that all the pictures used in the experiments had a high predictability and subjective familiarity. 60 pictures were finally used for target naming. The English and Chinese names of these 60 pictures had no difference in lexical frequency ($t = -1.42$, $p > 0.05$), all the 60 English or Chinese names of these pictures were nouns and had the same number of syllables. They all consisted of one syllable (CV or CVC). The rest 30 picture name were used as fillers.

After signing the consent form, all the participants completed a language proficiency questionnaire. They were seated in a soundproof chamber in front of a computer screen and were instructed to complete two experimental tasks in sequence, with a 20-minute interval between each. They were first required to complete the arrow-based spatial conflict task, and then, after a break of 20 minutes, they were instructed to finish a cued picture naming task, which consisted of a practice part and an experimental part.

With respect to the arrow-based spatial conflict task, participants are required to indicate the direction in which an arrow is pointing by pressing a key with their left or right index finger. Specifically, they press the "A" key with their left index finger when the arrow points left, and the "L" key with their right index finger when the arrow points right. In trials where the position of the arrow (left or right half of the screen) does not align with the direction the arrow is pointing, a conflict arises, leading to increased response times. The efficiency of inhibition is gauged by comparing response times in conflict trials with those in congruent trials. The version of the task used in the study consisted of 6 practice trials to familiarize participants with the task requirements and 120 critical trials, including 60 congruent trials where the arrow's position matched its direction and 60 incongruent trials where a conflict existed between the arrow's position and direction. This setup allows for the measurement of participants' inhibitory control by analyzing their response times in both conflict and congruent conditions.

In the Cued picture naming tasks conducted in this study, stimuli (pictures) were displayed at the center of a 19-inch computer screen with a resolution of 1024 × 768 pixels, viewed from a distance of approximately 80 cm. Each trial commenced with a 500 ms presentation of a cross, followed by a 500 ms blank screen interval. Subsequently, a picture stimulus appeared on the screen for 3000 ms, during which participants were tasked with naming the picture in either English or Chinese based on a language cue indicated by the background color—red for Chinese naming and blue for English naming. The background color cue was counterbalanced across all blocks. Participants were instructed to press the Enter key upon naming the pictures to proceed to the next trial. The Cued picture naming tasks were divided into four blocks. The first and second blocks involved non-language switching, where participants consistently named pictures in either English or Chinese, serving as controlled conditions. The third and fourth blocks introduced language switching scenarios, comprising two trial sequences: switch trials (alternating between English and Chinese naming) and repeat trials (repeating the same language for consecutive trials). The third block emphasized English dominance, while the fourth block emphasized Chinese dominance, with each block comprising 30 trials. In total, each participant completed 120 trials, and their picture naming responses across all blocks were recorded using Praat software on the computer.

In the experiment, all the sound records of the participants' naming pictures were analyzed using Praat, a professional phonetic analysis software developed by Boersma and Weenink in 2017. Praat is commonly used to measure various phonetic features of speech sounds, such as Voice Onset Time (VOT), which is the time duration between the release of a stop consonant and the onset of voicing (Boersma & Weenink, 2017). To measure the VOT of a phone, such as the /p/ sound in the word "pig," the following steps were followed: (1) Locate the point where the stop consonant is released and the onset of voicing begins. (2) Select the time span between these two points in the waveform. (3) Read the duration of this selection in milliseconds from the duration bar at the bottom of the editor window or obtain the selection length and read it from the information window. (4) Since the voicing onset follows the stop release, the VOT value is positive.

In the specific example provided, the VOT for the /p/ sound in "pig" was measured to be 76.63 milliseconds. This process illustrates how Praat can be utilized for precise phonetic analysis, allowing researchers to examine and quantify various acoustic properties of speech sounds accurately.

In addition to analyzing the participants' voice recordings and measuring the VOT of stops in English or Chinese during the cued picture naming task, 48 native speakers were recruited for further analysis. Among these speakers, 24 were native English speakers from the USA and 24 were native Chinese speakers. These participants were tasked with reading aloud all the

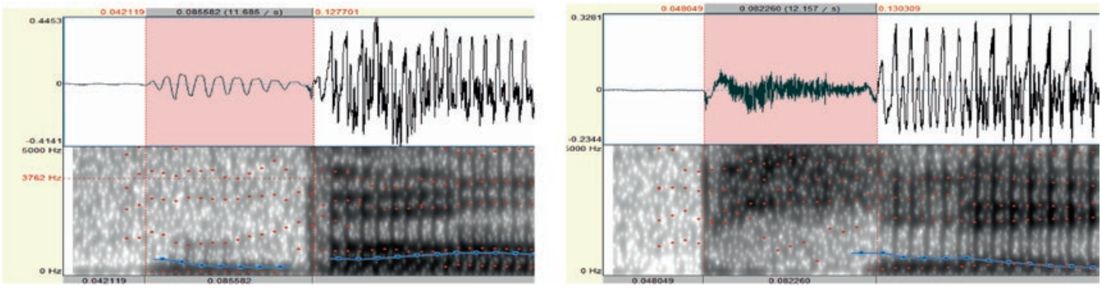

Figure 1. *VOT of English stops (/p/ /b/) in a word*

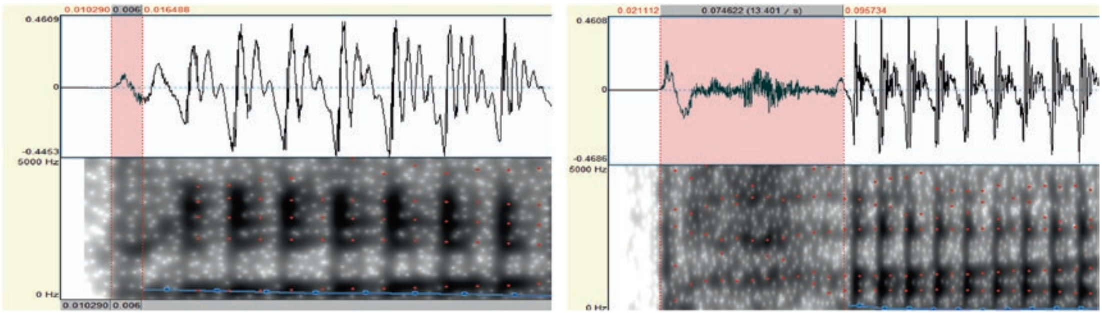

Figure 2. *VOT of Chinese stops (/p/ /p^h/) in a word*

target names of pictures used in the study. The voices of these target names were recorded and analyzed using the Praat software. Furthermore, the VOT of stops, including the /b/, /p/, /t/, /d/, /k/, and /g/ sounds in the target names, was measured as a baseline for comparison across native English and Chinese speakers. This comprehensive analysis aimed to capture and compare the phonetic characteristics and VOT values of stops in the names pronounced by native speakers of English and Chinese. By examining the VOT of stops in the target names spoken by native speakers of both languages, the research can gain insights into potential differences or similarities in phonetic realizations and articulatory patterns between English and Chinese speakers. This comparative analysis contributes to a better understanding of how VOT and stop consonant productions may vary across different language backgrounds.

Finally, this study was a mixed design of within-subjects and between-subjects: 2 (switch trial vs nonswitch trial) × 2 (L1-L2 switch vs L2-L1 switch) × 3 (high inhibition ability, mid inhibition ability, and low inhibition ability). In the cued picture naming task, the first and second blocks consisted of non-switch trials, while the third and fourth blocks were composed of switch trials. It should be mentioned that Chinese was actually a dominant language over the whole study. Regarding the switch sequence, two types of switch orders were implemented in the third and fourth blocks: switching from L1 to L2 or switching from L2 to L1. All the above variables were within subjects, while the participants' inhibition ability was a between-subject variable with three levels: high, mid, and low.

Data Analysis

All the data in the experiment were first collected in E-Prime by a computer and then were analyzed in SPSS 17.0. Some data were finally removed since they did not meet the experimental demand. In the cued picture naming task, all the participants' data were effective, while in the arrow-based spatial conflict task, only 31 participants' data were effective in the final data analysis because 13 participants' data were polluted. T-tests, Correlation and Repeated-Measures ANOVA were adopted to investigate the effects of inhibition ability on L2 leaners' VOT.

Results

To test the predictions outlined in the previous part that asymmetrical switch cost in terms of reaction time or VOT of stops would be found between different switch trials (L1-L2 vs L2-L1 switch) and a correlation would be found between inhibition ability and switch cost in terms of VOT if inhibition works in bilingual language processing, we analyzed switch cost pattern and the correlation between inhibition ability and switch cost in VOT. VOT in the first block (English without switch) and the second block (Chinese without switch) was adopted as a control condition, and the switch cost was obtained by subtracting the VOT of stops in the non-switch blocks from that in the switch blocks. The detailed results were presented as follows:

First came the switch cost in reaction time between two switch sequences (L1-L2 vs L2-L1).

Table 2. *Switch cost in reaction time between switch sequences (L1-L2 vs L2-L1)*

Switch sequence	Mean	N	SD	t	P
L1→L2	−57.45	40	463.45	−1.77	0.09
L2→L1	151.25	40	344.02		

Table 2 showed that in terms of reaction time, a marginal significant difference was obtained between L1-L2 switch cost and L2-L1 switch cost, $t = -1.77$, $P = 0.09$. The mean in L2-L1 switch was much larger than that of L1-L2 switch (151.25 > −57.45).

The second analyzed result falls on the variation of voice onset time of stops in English or Chinese. It is shown that the voice onset time of voiceless aspirated /p^h, / t^h/, and /k^h / (means = 87.92, 107.80, 112.16) in Chinese was much longer than that of voiceless unaspirated /p/, /t/, and /k/ (means = 12.69, 11.64, 29.00). On the whole, the voice onset time of Chinese stops did not have any significant difference across the conditions except for aspirated stop /t^h / and unaspirated /t/ that had significant differences between the non-switch condition and the switch condition (t = 2.288, 1.925, 6.097, 4.468, $Ps < 0.05$). In the same vein, nor did the voice onset time of English stops have any significant difference between the non-switch condition and the switch condition, except that the voice onset time of the aspirated stop /p/ in the switch condition differed significantly from that in the non-switch condition (t = 2.483, 2.378, $Ps < 0.05$). As compared with voice onset time of stops made by native speakers, that of all the stops made by Chinese-English bilinguals was quite different, t = −4.144, −3.342, 2.656, −4.224, −7.21; 2.853, −3.887, −7.394, −4.386, −4.065, −4.710, −3.511, −2.787; −2.325, −12.254, −6.380, −5.193, −52.111, −35.459, −29.009, −161.95, −56.156, −46.106, −72.734, −57.105, −59.616. $Ps < 0.05$.

The third analyzed results were concerned with the correlation between switch cost and inhibition ability. In the study, switch cost in VOT was obtained when VOT in the switch trials minuses VOT in the control condition (without any switch).

In Table 3, it was shown that the main effect of switch sequence (L1-L2 vs L2-L1) was significant, $F(1, 29) = 36.65$, $p < 0.001$, $\eta^2 = 0.56$. Table 3 showed an interactive effect of switch

Table 3. *ANOVA test of within-subjects effects of switch sequence and inhibition ability*

Source	df	MS	F	P.	η^2
Switch sequence					
Sphericity Assumed	1	2443.515	36.65	.00	.56
Greenhouse-Geisser	1.00	2443.515	36.65	.00	.56
Huynh-Feldt	1.00	2443.515	36.65	.00	.56
Lower-bound	1.00	2443.515	36.650	.000	.56
Switch sequence* Inhibition ability					
Sphericity Assumed	2	430.77	6.46	.005	.31
Greenhouse-Geisser	2.00	430.77	6.46	.005	.31
Huynh-Feldt	2.00	430.77	6.46	.005	.31
Lower-bound	2.00	430.77	6.46	.005	.31
Error(s)					
Sphericity Assumed	29	66.67			
Greenhouse-Geisser	29.00	66.67			
Huynh-Feldt	29.00	66.67			
Lower-bound	29.00	66.67			

sequence and inhibition ability on switch cost (in terms of VOT cost), $F (2, 29) = 6.46$, $P < 0.01$, $\eta^2 = 0.31$. This interactive effect was further analyzed in the following results. Switch cost in VOT had a significant difference between low and high or low and mid inhibition groups ($Ps < 0.05$); no significant differences were found between mid and high groups ($Ps > 0.05$). These differences were more significant in L2-L1 switch order than in L1-L2 switch order.

Table 4. *Correlation between inhibition ability and switch cost*

Items	1	2	3
1. VOT Cost (L1-L2)	1		
	32		
2. VOT Cost (L2-L1)	.33	1	
	.07		
	32	32	
3. Inhibition ability	.37*	.73**	1
	.04	.00	
	32	32	32

Table 4 showed a strong correlation between inhibition ability and VOT cost (L1-L2: $r = 0.37$; L2-L1: $r = 0.73$), particularly when switching to L1, a dominant language over all the conditions, $r = 0.73$, $P < 0.01$.

Discussion

The present study investigated whether inhibition can occur at the phonetic level when bilinguals switch from one language to the other. Three hypotheses were tested: (1) Reaction time in picture naming and VOT of stops produced by Chinese and English bilinguals would differ significantly between the switch condition and non-switch condition; (2) asymmetrical switch cost in VOT would be obtained between L1-L2 switch and L2-L1 switch; and (3) A strong correlation between inhibition ability and switch cost (VOT cost) could be found.

Based on Green's (1998) inhibitory control model in language switching paradigms, there is a typical pattern known as the switching cost, where bilinguals experience a larger cost when switching to their first language (L1) compared to switching to their second language (L2). This suggests that a greater cognitive effort may be required when switching back to L1, reflecting the inhibitory control processes involved in language switching. In the context of the study outlined, the hypotheses proposed are as follows: Hypotheses 1 and 2 predict that the VOT cost in L2-L1 switch trials will be significantly higher than that in L1-L2 switch trials for Chinese and English bilinguals when producing the names of pictures under different conditions. This aligns with the typical switching cost pattern observed in bilingual language production tasks.

The third hypothesis posits that individuals with greater inhibitory control abilities will demonstrate more efficient language switching, resulting in smaller switch costs. Specifically, individuals with better inhibitory control should exhibit smaller VOT switch costs for Chinese and English stops, with a stronger relationship observed when switching into L1. If these hypotheses are supported by the study's results, it would suggest that inhibitory control mechanisms play a role at the phonological level in bilingual speech production. This would provide valuable insights into the cognitive processes involved in language switching and production, highlighting the importance of inhibitory control in managing linguistic competition and facilitating efficient language switching.

The first hypothesis was partially verified. In the study, Chinese and English bilinguals were required to name pictures in the different conditions (switch conditions and non-switch conditions). Their reaction time in naming each picture and VOT of phonetic stops in the names they used were recorded and analyzed. Though the mean of reaction time in the switch condition was longer than that in the non-switch condition, the results showed no significant difference in reaction time or VOT between the switch condition and the non-switch condition. This might be attributed to the Chinese dominance over the whole experiment, as all the participants were late bilinguals who learned English at a later age. Their English proficiency was very low, which meant that little influence could be exerted on their Chinese. Conversely, the Chinese impact was apparent in that all the English stops VOT was much longer as compared with native English speakers. It seemed that all the participants transferred the Chinese phonetic features of stop categories into the English stops so that all English voiced unaspirated stops were pronounced in both switch conditions and non-switch conditions in a much longer VOT (larger than 0 ms) than the VOT (usually negative or short lag) made by the native speakers.

Regarding the second hypothesis, despite no findings of asymmetrical cost in reaction time, asymmetrical cost in VOT was observed between two different switch orders (L1-L2 vs. L2-L1) in the study. It meant that when switching into L1, the participants' Chinese stop VOT suffered more than their English did when switching into L2. This finding was quite contrary to the results that Linck (2008), Declerck et al. (2013), and Yang (2013) found in their studies.

Linck (2008) and Yang (2013) claimed that both RTs and accuracy data failed to show any asymmetrical switch cost in both the lexical decision task and the word naming task, and they assumed no evidence of L1 inhibition at the phonological level. Moreover, Declerck et al. (2013) also found no switch cost difference between words that consisted solely of phonemes that occur in both languages and words that contained language-specific phonemes. However, upon closer examination of their studies, it was found that Linck (2008) and Yang (2013) focused solely on testing reaction time (RT) and accuracy, rather than examining switch costs in specific phonemes. For example, in the lexical decision task, only a button-press button response was required while in the word naming task, the participants were only instructed to speak loudly “Yes” if it was a word or “No” if it was not a word. It may not require any processing at the level of phonology. The distinction between the phonological and lexical levels is crucial in language processing research, as these levels play distinct roles in speech production and comprehension. It is crucial to carefully distinguish between these levels to accurately interpret the findings of studies examining language switching and its impact on phonetic and phonological processes in bilingual individuals. Recent studies, such as those by Olson (2013) and Goldrick et al. (2014), have shed light on the impact of language switching on accents and phonetic production, providing valuable insights into the interaction between the phonetic systems of bilingual speakers. In Olson’s study (2013), the investigation into the effect of language switching on phonetic production revealed asymmetrical effects on voice onset time, with the dominant language being significantly influenced by the act of switching languages. This asymmetry suggests that language switching can have differential effects on phonetic features in bilingual speech production tasks. By demonstrating the impact of language switching on phonetic production and highlighting the asymmetrical effects on phonetic features, studies like Olson (2013) contribute to our understanding of how bilingual individuals manage and coordinate their two phonetic systems during language switching tasks. The findings suggest that inhibitory control mechanisms may operate at the phonological level, influencing phonetic production and the interaction between a bilingual’s linguistic systems. These studies provide valuable insights into the complex interplay between language switching, phonetic production, and inhibitory control mechanisms in bilingual speech processing, furthering our understanding of the underlying cognitive processes involved in bilingual language use.

Regarding the third hypothesis, a direct link was found between individuals’ inhibitory ability and switch cost. A series of results in the experiment provided clear support for this hypothesis. It was found that when switching into L1, the switch cost in voice onset time strongly correlated with the participants’ inhibition ability. When switching into L2, the switch cost of voice also correlated with inhibition ability, but not as strongly as switching into L1. The finding was similar to Linck et al (2011)’s research, in which they found that better inhibitory control was related to reduced switch costs, but only when switching into or out of the more dominant L1 and concluded that there was a direct link between inhibitory control abilities and language switching capabilities. In other words, the participants with low inhibitory abilities suffered more in switching trials. When switching to L1 (Chinese), the onset time of their voices in Chinese stops was more influenced, and their phonological production in naming pictures might not be native-like. Just like what Lev-Ari and Peperkamp (2013) found, bilinguals with lower inhibitory skills exhibited greater influence of the second language on the first, which led to non-native perception and production in bilinguals’ native language. Since Linck et al. (2011)’s study provided a more fine-grained analysis of the relationship between language control and executive functions by linking inhibitory abilities to individual differences in online lexical access during speech production (Linck et al., 2011), the current study provided certain evidence for the strong link between inhibitory abilities and individual differences in online phonological processing during speech production.

The aforementioned results and analyses demonstrated that asymmetrical switch cost in voice onset time of stops between language switching was found, and better inhibitory abilities could predict better performance in switch cost and suffer less. This was the very prediction that Green (1998)'s IC model made, which assumed that there were asymmetrical switch costs, with larger switch costs into the more dominant language, which had been proved by ample research (Meuter & Allport, 1999; Philipp, Gade & Koch, 2007; Fan et al., 2013; Yang, 2016). The bilingual language production model outlined in the context of this study suggests that bilingual individuals have a shared pool of phonological features that are utilized across both languages. This model posits that when producing spoken words, the phonology of the word is specified, and that common phonological elements in each language can activate similar or identical representations. As a result, bilinguals, especially those who are proficient, have multiple word options available for each concept, leading to potential competition and interference among activated phonological information during language production. A cognitive mechanism is posited to modulate the selection process and control performance to address this competition and interference. The study identified inhibition as the cognitive mechanism that operates locally at specific points to regulate the competition and influence the Voice Onset Time (VOT) of stops at the phonological level. This suggests that inhibitory control is crucial in managing linguistic competition and selection processes during bilingual language production tasks.

Conclusion

This study investigated whether inhibitory control plays a role at the phonetic level in speech production during language switching between two languages. This investigation was carried out in two-fold parts. Firstly, the study examined the impact of language switching on specific phonetic features by utilizing a cued picture naming task. Secondly, the study sought to establish a potential correlation between individuals' inhibition abilities and the switch cost at the phonetic level, specifically focusing on the VOT of stops. The results yielded two principal findings: 1) when switching into L1, a dominant language of Chinese-English bilinguals, asymmetrical switch cost in voice onset time of stops was found, which was believed to be a strong index of inhibitory control that worked at a specific point; 2) a strong correlation was obtained between inhibition abilities and switch cost (VOT cost), better inhibition abilities led to a less cost of VOT. These findings suggest that in Chinese and English bilinguals' picture naming, inhibitory control might be needed to suppress the influence of one language's phonetic system on the other language, and better inhibitory control abilities might lead to less influence from the other language in language switching. The results would contribute to our understanding of the cognitive processes involved in language switching and phonetic production in bilingual individuals, particularly, providing a support for Green (1998)'s Inhibitory Control Model from a phonological level. They can also inform pedagogical practices for bilingual education by emphasizing the need to train inhibitory control to improve phonological accuracy during language switch and reduce cross-linguistic interference. Overall, this work shifts the discourse from purely lexical bilingual processing to integrated cognitive phonological dynamics, urging future studies to consider inhibitory control as a core component in both theoretical models of bilingualism and evidence-based language pedagogy.

Acknowledgments

This study was funded by JiangSu Social Science (Grant 23YYB004).

References

Antoniou, M., Best, C. T., Tyler, M.D., & Kroos, C. (2011). Interlanguage interference in VOT production by L2 dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics*, 39, 58–570.

Antoniou, M., Tyler, M. D., & Best, C. T. (2012). Two ways to listen: Do L2 dominant bilinguals perceive stop voicing according to language mode? *Journal of Phonetics, 40*, 582–594.

Bajo, M.T., Gomez-Ariza, C.J., & Marful, A. (2021). Inhibitory control of information in memory across domains. *Current Directions in Psychological Science,* 30(5), 444–453.

Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism*, *9*, 103–119.

Boersma, P., & Weenink, D. (2017). Praat: Doing phonetics by computer [Computer program]. Version 6.0.25, retrieved 11 February 2017 from http://www.praat.org/.

Broersma, M. (2011). Triggered code-switching: Evidence from picture naming experiments. In M.S. Schmid & W. Lowie (Eds), *Modeling bilingualism: From structure to chaos: In honor of Kees de Bot* (pp. 37–58). Amsterdam, The Netherlands: Benjamins.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, 27, 341–350.

Coderre, E., van Heuven, W., & Conklin, K. (2010). *Lexical access and executive control in monolinguals and bilinguals*. Poster presented at the Architectures and Mechanisms of Language Processing (AMLaP) Conference, York, UK.

Dai, Z. Z. (2012). A study of VOT of stops produced by Chinese learners of English. *Journal of Xi An Electronic Science and Technology University (Social Science)*, *22*(6), 102–106.

Darcy, I., Mora, J. C., & Daidone, D. (2016). The role of inhibitory control in second language phonological processing. *Language learning,* 66(4), 741–773.

Declerck, M., & Koch, I. (2023). The concept of inhibition in bilingual control. *Psychological Review*, *130*(4), 953–976.

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

Declerck, M., Koch, I., & Philipp, A.M. (2015).The minimum requirements of language control: Evidence from sequential predictability effects in language switching. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *41*, 377–394.

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

Declerck, M., Philipp, A.M., & Koch, I. (2013). Bilingual control: Sequential memory in language switching. *Journal of Experimental Psychology: learning, Memory and Cognition*, 39, 1793–1806.

Dong, Y. P. & Li, P. (2020). Interpreting: A window into bilingual processing. *Bilingualism: Language and Cognition*, *23*(4), 1–3.

Fan, L., & Li, S. S. (2013). Inhibitory processing of code-switch in Chinese-English-Japanese trilinguals production-based on stimulus and response design, *Journal of Foreign Language Teaching and Research*, *45*(1), 58–68.

Filippi, R., Karaminis, T., & Thomas, M. S. C. (2014). Bilingual language switching in production: Empirical and computational studies. *Bilingualism: Language and Cognition*, *17*, 294–315.

Fowler, C.A., Scramko, V., Ostry, D. J., Rowland, S. A., & Halle, P. A. (2008). Cross language phonetic influence on the speech of French-English bilinguals. *Journal of phonetics*, *36*, 649–663.

Goldrick, M., Runnqvist, E., & Costa, A. (2014). Language switching makes pronunciation less nativelike. *Psychological Science*, *25*, 1031–1036.

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

Grosjean, F., & Miller, J. L. (1994). Going in and out languages: An example of bilingual flexibility. *Psychological Science*, *5*, 201–209.

Keating, A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, *60*(2), 286–318.

Kerz, E., & Wiechmann, D. (2017). Individual differences in L2 processing of multi-word phrases: Effects of working memory and personality. In: Mitkov, R. (eds) Computational and corpus-based phraseology (pp. 306–321). *EUROPHRAS 2017. Lecture notes in computer science*, Vol 10596. Springer, Cham.

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences*, February, *22*(2), 154–169.

Lev-Ari, S., & Peperkamp, S. (2013). Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *Journal of Phonetics, 41*, 320–331.

Lev-Ari, S., & Peperkamp, S. (2014). The influence of inhibitory skill on phonological representations in production and perception. *Journal of Phonetics,* 47, 36–46.

Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. (2007). Inhibiting your native language: The role of retrieval induced forgetting during second language acquisition. *Psychological Science*, *18*(1), 29–34.

Li, L., & Dong, Y.F. (2014). Perception cues of Chinese stops: Differences between the native speakers and second language learners. *Psychological Research*, 7(1), 41-45.

Li, H. L., Goldin, P., & Siegle, G. (2021). Neuroscience for clinicians: translational clinical neuroscience to inspire clinical practice and research. Reference Module in Neuroscience and Bio-behavioral Psychology.

Linck, J. A. (2008). *The role of inhibition in the control of bilingual speech production*. Doctoral dissertation, The Pennsylvania State University.

Linck, J. A., Schwieter, J. W., & Sunderman, G. (2011). Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism: Language and Cognition*, *15*, 651662.

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops. *Word*, 20, 284-422.

Liu, X. J. (2024). Review of cognitive individual differences in second language acquisition. *Australian Journal of Applied Linguistics*, 7(1), 1–5.

Marian,V., Blumenfeld, K., Mizrahi, E., Kania, U., & Cordes, K. (2012). Multilingual stroop performance: Effects of trilingualism and proficiency on inhibitory control. *International Journal of Multilingualism*, 1, 1–23.

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

Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition, 21*, 81–108.

Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*, *41*, 407–420.

Philipp, A., 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.

Ran, Q. B., & Shi, F. (2007). An analysis of VOT of Plosives in Chinese mandarin monosyllables, *Journal of Nan Kan Linguistics*, *2*, 23–30.

Schwieter, J. W., & Sunderman, G. (2008). Language switching in bilingual speech production: In search of the language specific selection mechanism. *The Mental Lexicon*, *3*, 214–238.

Schwieter, J., & Wen, Z. (2022). The Cambridge handbook of working memory and language. CUP.

Simon, J. R & Rudell, A. P. (1967). Auditory S-R compatibility: The effects of an irrelevant cue on information processing. *Journal of Applied Psychology*, 51, 300–304.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, *30*(4), 510–532.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology General*, *18*(6), 643–662.

Thomas, M. S. C., & Allport, D. A. (2000). Switching costs in bilingual visual word recognition. *Journal of Memory and Language*, *43*(1), 44–66.

Topal, I. H. (2024). Turkish EFL learners' knowledge, perception, and production of suprafixes as part of lexical competence. *Australian Journal of Applied Linguistics*, 7(3), 1530.

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

Wang, M. L. (2009). An analysis of pronunciation of stops in English words of Chinese learners of English, *Journal of Modern Foreign Languages*, *32*(2), 186–194.

Wen, Z. S & Li, S. F. (2019). Working memory in L2 learning and processing, In Schwieter, W., & Benati, A. (eds), *The Cambridge handbook of language learning* (pp. 365–389), Cambridge University Press.
Yang, W. Q. (2013). *Language control in bilingual production—a new study of the role of inhibition*, Unpublished Doctoral Degree paper, FuJian Normal University.
Yang, W. Q., & Qin, Y. X. (2016). Inhibition in language control in Bilingual language production, *Journal of Modern Foreign Languages*, *3*, 381–389.
Zheng, X. R., & Li, Y. H. (2007). A contrastive analysis of VOT of stops between English and Chinese and a study of acquisition of English stops of Chinese learners, *Journal of ChangChun Normal University (Social Science)*, *26*(1), 92–95.