Bilingualism: Language and Cognition

cambridge.org/bil

Language switch costs in sentence comprehension between Chinese and English: Evidence from self-paced reading

Mengyan Zhu[1,2], Patrick Sturt[2] and Markus Damian[1]

[1]School of Psychological Science, Department of Life Sciences, University of Bristol, Bristol, BS8 1QU, UK and [2]School of Philosophy, Psychology and Language Sciences, The University of Edinburgh, Edinburgh, EH8 9JZ, UK

Research Article

Cite this article: Zhu, M., Sturt, P. and Damian, M. (2025). Language switch costs in sentence comprehension between Chinese and English: Evidence from self-paced reading. *Bilingualism: Language and Cognition*, 1–15. https://doi.org/10.1017/S1366728925000100

Received: 02 June 2024
Revised: 08 January 2025
Accepted: 09 January 2025

Keywords:
language switching; language comprehension; switch cost; inhibitory control; bilingualism

Corresponding author:
Mengyan Zhu;
Email: Mengyan.zhu@bristol.ac.uk

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

Abstract

Effects of language switching in bilinguals have been extensively investigated, but the majority of studies have focused on switching in language production. Here we explored intrasentential switching between Chinese and English, employing a self-paced reading paradigm, with Chinese/English using radically different orthographic systems. In addition, we investigated whether L2 (English) proficiency influences switch costs. Results revealed that switch costs emerged only when switching into L1 (Chinese); by contrast, when switching into L2, a less reliable facilitatory effect was observed. L2 proficiency affected reading speed for English stimuli, but it did not directly modulate switch costs. Moreover, we have integrated various findings from previous research and identified that the use of different comparison patterns is a major contributing factor to the inconsistency in results among prior studies. We suggest that in cross-script language switching, switch costs stem from a general cognitive control mechanism rather than from activation within the bilingual mental lexicon.

Highlights

- We explored switching between Chinese and English in sentence reading
- Self-paced reading task was used to capture the effects of the language switch
- Switching into L1 (Chinese) but not into L2 (English) incurred a cost
- L2 proficiency was not the main determinant of switch cost
- Switch costs stem from cognitive control rather than from lexical co-activation

1. Introduction

Language switching is a prevalent occurrence in societies with multiple languages (Wang, 2015) and bilingual speakers are often forced to, or voluntarily choose to, switch between their languages, both in language production (for instance, when intermixing words or phrases from multiple languages) and in comprehension (when trying to understand an utterance that contains portions of multiple languages). Numerous experimental studies have investigated the consequences of language switching (see Declerck & Philipp, 2015, for review). The general observation is that switching languages incurs a cost. For instance, in language production, when individuals name pictures and are on each trial cued as to which language to use for the response, responses are slower when preceded by a trial that involves a different language than on a preceding trial that involved the same language (e.g., see Johns & Steuck, 2021; Meuter & Allport, 1999; Verhoef et al., 2009; but see Blanco-Elorrieta & Pylkkänen, 2018). Switch costs in language production are also often found to be asymmetric, with larger switch costs when switching into L1 than into L2 (but this conclusion is not universally accepted; see Gade et al., 2021, Goldrick & Gollan, 2023, for discussion). Compared to language switching in production, fewer studies exist on the consequences of switching in language comprehension. In single-word processing such as in lexical decision, semantic categorization or masked priming tasks, the evidence for language switch costs is not particularly clear, as will be reviewed below. *Intrasentential* switching (Poplack, 1980) in bilingual reading and its consequences are a particularly understudied area. This topic is the focus of the work presented below.

1.1. Language switching

The inhibitory control model (ICM; Green, 1998) provides an influential account of how switch costs may arise in language tasks and why switch costs may be asymmetrically dependent on language dominance. According to this model, a bilingual individual will co-activate the representations of their multiple languages. Each language task activates its own ‘task schema’ and a switch in response languages necessitates a schema change resulting in slowed latencies.

CAMBRIDGE UNIVERSITY PRESS

Further, activation of one schema leads to the reactive inhibition of the other, hence avoiding interference between lemmas in different languages (Bialystok et al., 2009; Green & Abutalebi, 2013; Hilchey & Klein, 2011). The asymmetry in switch cost is explained by the need to prevent conflict between the target and non-target word lexicons. This entails a greater cognitive effort to suppress the dominant language (L1) when processing the non-dominant language, resulting in increased effort to return to L1 afterward. Conversely, inhibiting the non-dominant language (L2) during dominant language processing is unnecessary, making it easier to switch back to L2 than to L1. This principle of co-activation followed by reactive inhibition applies not only to language production but also to receptive language activities such as reading or listening (Green, 1998, p. 74).

It is, however, not clear whether language production and reception are directly comparable regarding task demands and requirements (Wang, 2015). In production, the bilingual speaker ultimately chooses the target language, whereas processing in receptive activities is more 'bottom-up' and conceivably requires less cognitive control. Given that the goal of language comprehension is to decode a linguistic signal, it makes sense for readers/listeners to apply the entirety of their linguistic knowledge to the problem. Indeed, an extensive literature, mainly centred on the processing of single words, suggests that bilingual readers co-activate across their multiple representational systems. For instance, in a series of lexical decision and progressive demasking experiments, van Heuven et al. (1998) found cross-language neighbourhood effects (i.e., increasing the number of Dutch orthographic neighbours slowed down response latencies for English target words). At the same time, it is plausible (but not self-evident) that readers/listeners would inhibit co-activated words from the non-target language.

The prominent Bilingual Interactive Activation (BIA) model (Grainger & Dijkstra, 1992; Grainger, 1993; van Heuven et al., 1998) offers a theoretical framework for understanding cross-language co-activation and resolution of competition in response to written words. This framework assumes that a word activates within- and cross-language orthographic and phonological neighbours, as well as triggers the so-called 'language nodes', which ultimately determine the target language. Activated language nodes reactively inhibit words from the non-target language, with the strength of inhibition dependent on the language nodes' activation. This assumption predicts an asymmetry in switch cost dependent on language dominance but in the opposite direction to that derived from the ICM: when processing L1, the L1 language node is likely to be more strongly activated than the L2 language node is when processing L2; hence, inhibition of L2 words due to activation of the L1 language node is generally stronger than vice versa (Grainger et al., 2010). Therefore, switch costs should be larger when switching into L2 than into L1. However, in the subsequent, revised BIA+ model (Dijkstra & van Heuven, 2002), the principle of reactive inhibition of non-target language representations has been abandoned. Instead, activation from language processing is forwarded to a task/decision system where a decision is being made based on relevant task demands. Therefore, two systems are involved in the BIA+ model: word identification and task/decision systems (see more details in the Discussion).

Although it is well documented that bilingual readers generally co-activate their multiple linguistic systems, evidence mainly comes from studies using combinations of languages with similar orthographic scripts. When switching between alphabetically spelled words, on a given trial, readers may receive no cues as to which language they are dealing with, or cues may come from language-specific letter properties. However, when languages with very different orthographies are intermixed (e.g., Chinese and English), then each word will clearly announce its language membership at the orthographic level. In cases such as these, it is less clear whether bilingual readers co-activate multiple languages. According to Dijkstra and van Heuven (2002), '…no orthographically similar word candidates can be activated across language pairs that do not share orthography at all (e.g., Chinese and English), even though effects of phonological similarity might still occur for such language pairs. In other words, when particular input aspects are language specific, we will (of course) find evidence of language-specific access (e.g., Chinese orthography will not induce much Latin letter activation)' (p. 183). This makes the exploration of language switching between combinations such as Chinese/English particularly interesting: if a switch cost were to be found, it would be unlikely to take place within the language system itself, but rather it would arise at the task/decision system.

Are receptive activities such as those involved in visual word recognition subject to switch costs and, if so, are these asymmetric? This empirical question is surprisingly difficult to resolve, with a number of reports of switch costs in receptive tasks such as lexical decisions and semantic categorizations (von Studnitz & Green, 1997; Thomas & Allport, 2000), but also prominent recent null findings (Declerck, Koch, Dunabeitia et al., 2019). A difficulty is that typical single-word processing tasks require the classification of stimuli into one of two responses and a corresponding keypress response. This introduces the potentially relevant additional variable of whether or not on a given trial, the *response* is the same or different, relative to the previous trial. In the field of task (rather than language) switching, it is well known that in switch tasks, repeated responses on consecutive trials lead to faster responses on non-switch trials but to slower responses on switch trials (e.g., Rogers & Monsell, 1995). Hence, switch costs interact with response repetition in a complex way that allows for various theoretical interpretations (see Koch et al., 2023). Any study on language switching, which requires binary responses but does not take into account the response factor, will therefore be difficult to interpret. This is, for instance, the case for lexical decision studies reported by Aparicio and Lavaur (2014), Koeth (2012), Ong et al. (2019) and von Studnitz and Green (1997). Recently, Struck and Jiang (2022) explicitly modelled switch costs and response repetition in an experiment in which participants carried out lexical decisions on either English or Chinese words. Results showed a complex pattern with a switch benefit on response-change trials but a switch cost on response-repeat trials in L1; however, it showed a switch cost on both response-change and response-repeat trials in L2. Further research is needed to determine the effect of language switching in bilingual single-word processing.

1.2. Intrasentential language switching

The current study investigates whether intrasentential language switching incurs a cost, whether this cost is asymmetric depending on switch direction and if it is modulated by language dominance. Prior research (e.g., Beatty-Martínez et al., 2021; Adler et al., 2020) has examined the effects of language switching in both spoken and written sentences. For instance, Altarriba et al. (1996) studied Spanish–English bilinguals' responses to embedded target words differing in frequency and sentence constraint. Eye movements in natural reading and naming of the target word following rapid serial visual presentation were measured. However, this study

https://doi.org/10.1017/S1366728925000100 Published online by Cambridge University Press

focused on the role of lexical and conceptual constraints and did not primarily explore language switch costs. Adler et al. (2020) asked Spanish–English bilinguals to read sentences that were presented word-by-word in a self-paced reading task, with sentences either presented in English, Spanish or beginning in Spanish and ending in English. Average word reading times appeared to be affected by switching in English but not in Spanish; however, the primary aim of this study was to explore potential consequences of a language switch on interleaved Flanker task performance and so self-paced reading times were not analysed by word position. Salig et al. (2023) extended this by exploring the frequency effects of switching patterns on code switching within verb and noun phrases for Spanish–English bilinguals.

A straightforward approach to study intrasentential switch costs is to present bilingual participants with sentences in L1, L2 or with a language switch, then analyse word reading times around the switch point. This can reveal if language switching slows reading and whether effects vary with L1/L2 dominance. Bultena et al. (2015) reported a study in which Dutch–English participants engaged in a self-paced reading task, with the critical switch position within target sentences identified after a cognate/non-cognate verb. They found a cost when switching into L2 but not when switching into L1. Beyond the observed asymmetry, they also noted the influence of L2 proficiency on the magnitude of switch costs: lower proficiency readers spent more time reading switched sentences than their higher proficiency counterparts. The researchers argued that these findings reflect the activation strength in the L1 and L2 lexicons, emphasizing the role of relative lexical activation levels influenced by language proficiency (Alvarez et al., 2003). They suggested that an inhibitory control mechanism that operates in a top-down fashion is less likely to impact reading that operates in a bottom-up manner, at least in the initial processing stages. Activating L2 demands more effort than L1, resulting in a longer switching time, and higher L2 proficiency readers possess a stronger L2 lexicon, making it easier to activate compared to readers with lower L2 proficiency.

Few studies have explored the intrasentential switching between language combinations that involve different orthographic systems. For combinations such as Chinese/English, differences exist not only regarding their orthographies but also in sentence structure and syntax. For instance, relative clauses are less common in Chinese, with alternative structures conveying relationships between ideas. While Chinese employs fewer conjunctions, English utilizes a variety of conjunctions to connect clauses and express different relationships (Liao & Wang, 2015). The distinct syntactic differences between the languages pose potential challenges for cross-script language-switching studies. Recently, Hu and Zhao (2023) employed a self-paced reading task and an acceptability judgment task to investigate whether code-switching during syntactic processing in Chinese–English dual languages incurs a cost. They manipulated the word order of the embedded language to be consistent or inconsistent with that of the matrix language and took their results to argue that syntactic processing is a source of switch costs in sentence processing.

A further study that investigated switch costs with the Chinese/English language combination in a so-called 'maze task' was reported by Wang (2015). In the maze task (e.g., Forster et al., 2009), participants are presented first with a single word and then subsequently with a series of two-word alternatives to choose from, only one of which is grammatically acceptable as a sentence continuation. This task forces participants into a strictly incremental processing pattern: only after an explicit choice has been made in a given step of the maze is sentence continuation possible. By using simple and short sentence structures with consistent word order across all conditions, there are no significant syntactic differences. The switched trials followed the 'English–Chinese–English' language sequence, while the non-switched trials adhered to the 'English–English–English' pattern. For instance, a non-switched sentence is 'I polished my shoe yesterday', and a switched sentence is 'I polished my 鞋 yesterday' (鞋 means 'shoe' in Chinese). The comparisons were conducted in word region 2 (between 'shoe' and '鞋') and word region 3 (between two 'yesterdays' from different conditions).

In Wang's (2015) study, there were 50 English–Chinese bilinguals, split by English or Chinese dominance. Beyond the main effects of language dominance and switch cost, an interaction of language dominance and switch costs was found in word region 2 (comparing 'shoe' and '鞋'), but not in word region 3 (comparing two 'yesterdays'). The author suggested two mechanisms for switch costs in comprehension based on these findings: (1) an inhibitory control mechanism, typically operating for word region 3, where compared candidates are the same and there is no effect of language dominance on switch mode, indicating a non-lexicon-based mechanism, and (2) a lexical activation mechanism in word region 2, where language dominance influences the switch cost pattern. Specifically, Chinese-dominant readers read the word region 2 of switched sentences faster than non-switched sentences. By contrast, English-dominant readers read this word region slower in switched sentences than in non-switched sentences. This study concluded that inhibitory control is at least one of the mechanisms underlying switch costs in comprehension; it is noteworthy that this inference is at odds with the conclusions drawn from Bultena et al.'s (2015) study reviewed above.

Overall, evidence regarding intrasentential language switching is scarce and the results are somewhat inconsistent. The way in which conditions of switching and non-switching are compared in previous studies is also not consistent and this may contribute to the diverse pattern of results. Consider an intrasentential language-switching study schematized in Table 1. Words are presented either in the participants' native language (L1) or in their non-native language (L2). At some point in the sentence, the language either does or does not change, with reading times at or following the switch point of major interest because they might reveal the effects of language switching when switch and non-switch reading times are compared. The combination of two languages and switch/non-switch generates four sentence types: those who maintain the same language throughout (L1–L1 and L2–L2) and those in which a language switch occurs (L2–L1 and L1–L2).

The complexity arises from how switch costs are computed for L1 and L2 separately. Table 2 shows two possible comparison patterns. Comparison 1 keeps the pre-switch language constant while varying the switch/post-switch language; hence, to identify

Table 1. Structure of intrasentential language switching experiments

Language		
Pre-switch position	Switch/Post-switch position	Switch
L1	L1	No
L2	L2	No
L2	L1	Yes
L1	L2	Yes

https://doi.org/10.1017/S1366728925000100 Published online by Cambridge University Press

Table 2. Possible comparison patterns in language switching studies to capture effects of switching

	Comparison 1		Comparison 2	
	Switch	Non-switch	Switch	Non-switch
Switching to L1	L2–L1	L2–L2	L2–L1	L1–L1
Switching to L2	L1–L2	L1–L1	L1–L2	L2–L2

the effects of switching from L1 to L2, L1–L1 is compared to L1–L2, and to capture the effects of switching from L2 to L1, L2–L2 is compared to L2–L1. Comparison 2 keeps language at the critical (switch/post-switch) portion of the sentence constant but varies the language preceding the switch position. Hence, to identify the effects of switching to L1, L1–L1 is compared to L2–L1, and to reveal effects of switching to L2, L2–L2 is compared to L1–L2. It is worth noting that outside the intrasentential switch literature, comparison 2 is the method universally used in studies of language switching. For instance, in experiments on language switching in picture naming (e.g., Meuter & Allport, 1999), on a critical trial N, the language is constant (either L1 or L2) but the language relative to the previous N-1 trial is either switched or not.

Each comparison method has advantages and disadvantages. Comparison 1 creates a common baseline by keeping the language used in the first half of each sentence constant. Because reading times at or after the switch position are likely influenced by the words that preceded them and this part is held constant, it can be argued that reading times at or following the switch position will then reveal the true switch effects for each language. However, a disadvantage is that when analysing the critical portions of the sentences, this method compares reading times across L1 and L2. These are likely to differ in the first place because native reading is likely faster than non-native reading. Comparison 2 can be argued to start from different baselines, but at the critical sentence positions, languages are the same across switch and non-switch conditions, so the comparison is like-for-like.

Bultena et al. (2015) calculated switching effects using comparison 1 in Table 2. This approach also aligns with the word region 2 in Wang's study (2015). As highlighted above, doing so creates a common baseline but the critical switch versus non-switch comparisons in the second half of the sentence involves reading in different languages. Bultena et al. noted this aspect (see Footnote on p. 462) and argued that the approach of establishing a common baseline in the first half of a sentence, followed by a switch or non-switch into another language, is preferable. As summarized above, using this method they found a large cost when switching into L2 and a subtle switch benefit in L1. However, they reported that when they reanalysed their data according to the alternative method (comparison 2 in Table 2), then a large switch cost was found in L1 and a smaller one in L2. By contrast, Wang (2015) at the critical region 3 used comparison pattern 2, and switch costs to L1 are presumably revealed.

When surveying the literature on switch costs, we concluded that comparison 2 in Table 2 captures the idea of language switching more appropriately. Consequently, the data reported below were analysed according to this approach. Further insights into the effect of the comparison pattern are discussed in the Discussion section, while results analysed based on comparison pattern 1 are presented in the Appendix.

1.3. The present study

Below we report the results of a study, which explored intrasentential switching between Chinese and English in a self-paced reading task. We aimed to address two primary objectives: (1) to investigate switch effects in sentence comprehension in cross-script language combinations; (2) to assess whether the mechanism of switching relies on the relative activation level of the lexicon. Contrary to Hu and Zhao (2023) who had also explored Chinese/English language switching, we chose target sentences with simple sentence structures that had word-for-word matching translation equivalents (as had been the case in Bultena et al., 2015, for their Dutch–English sentences). As outlined above, we captured the effects of switching via a comparison pattern in which pre-switch position languages were different but switch and post-switch position languages were identical (i.e., comparison pattern 2 in Table 2). To fully assess the role that different comparison patterns may play in the mixed results from previous studies, we also conducted the analysis using comparison pattern 1. The detailed results are presented in the Appendix, and the results presented in the main text are based on the comparison pattern 2.

We further explored whether L2 proficiency influences the switch cost magnitude. Specifically, if high L2 proficiency readers demonstrate greater ease and speed in reading sentences switching to L2, this would suggest that the language-switching mechanism is lexicon-based and influenced by relative lexical activation strength. If there is no effect of L2 proficiency, it suggests that the relative activation of the lexicon is not the primary factor driving language switching between Chinese and English. This would imply that non-linguistic factors may play a more significant role (see the Discussion section for a more detailed explanation). Regarding inhibitory control, the specific switch pattern, such as the difficulty in switching into L1 compared to switching into L2, could be accounted for via the assumption of inhibitory control processes. Given the mixed results in previous research on switch costs in comprehension, predicting the switch pattern remains challenging. Even if we observed an asymmetric switching cost, with greater costs when switching into L1, this would not necessarily imply a role of inhibition. However, it is necessary to clarify the pattern of switching costs in language comprehension and this is also the main goal of the present study. Because assessing L2 proficiency through self-questionnaires may raise concerns about subjectivity (Wen & van Heuven, 2017), we employed a standardized English vocabulary test (LexTALE; Lemhöfer & Broersma, 2012) to quantify participants' English proficiency more objectively. Doing so allowed us to gain a clearer understanding of the potential role of L2 proficiency in influencing switch costs.

2. Methods

2.1. Participants

Sixty-one Chinese–English bilinguals (17 males and 44 females) aged between 18 and 30 years old (M = 22, SD = 3) were drawn from the participant pools of Shaanxi Normal University, China, and the University of Edinburgh, UK, and were paid 15 RMB[1]. All participants were Chinese native speakers and had learned English as a second language at an average age of 9 years

[1] Participants were recruited from two locations to ensure a range of English proficiency levels essential for the study. Some participants were Chinese students at the University of Edinburgh who had been in the United Kingdom for less than a year, likely with a higher English proficiency due to passing a language test for study purposes. Importantly, all participants were native Chinese speakers.

https://doi.org/10.1017/S1366728925000100 Published online by Cambridge University Press

(SD = 3). All participants had normal or corrected-to-normal vision and indicated no history of mental diseases. In the final data analysis, data from 53 participants were included (see the deletion criteria in the section Analysis). The experiment was approved by the PPLS Ethics Committee at the University of Edinburgh.

2.2. Materials and design

Materials consisted of 60 experimental sentences and 60 filler sentences, preceded in the self-paced reading task by seven practice sentences. Each experimental sentence was seven words long and, as mentioned earlier, the syntax of English and Chinese translation equivalents was consistent, ensuring that the word order remained unchanged across all conditions. Sentences were either fully presented in English, fully in Chinese, or the language switched from Chinese to English or vice versa. For the sentences with language switch, the switch position was always located at the fourth-word position (WP4)[2]. For example, an English sentence is '*Many kids like **this** interesting online game*'; its Chinese equivalent is '许多孩子喜欢这款有趣的线上游戏' (switch position is bolded for illustration only). Sentences were rotated across the four conditions (i.e., English–English, Chinese–Chinese, English–Chinese and Chinese–English) based on a Latin Square design, such that each sentence structure was presented only once to each participant. Filler sentences were included to introduce some variability regarding sentence length and structure, and switch position. These had an average number of 5.7 word positions (ranging from 4 to 8) and the switch positions varied from WP2 to WP5. In the experiment, a fresh random sequence of sentences was presented for each participant.

To ensure that participants properly processed the sentences, comprehension questions were constructed for half of the experimental and filler sentences, with participants instructed to answer 'yes' or 'no' by mouse click. The language used in the comprehension question was always the same as that in the second half of a corresponding sentence. As all the sentences were presented randomly, comprehension questions also appeared randomly.

2.3. Procedure

The study was conducted online using the Ibex Farm server (Drummond et al., 2016). Before the self-paced reading experiment, participants carried out the LexTALE test (Lemhöfer & Broersma, 2012) to assess their English (L2) proficiency. In this test, 60 letter strings (40 words and 20 pseudowords randomly intermixed, plus three initial practice strings were presented at the centre of the screen and participants judged whether each string was an English word or not, by selecting 'Word' or 'Non-word' with the computer mouse, without feedback. Then, the instructions for the self-paced reading task were given in Chinese, encouraging participants to read at a normal speed in word-by-word procedure and to answer yes/no comprehension questions following half of the sentences using the computer mouse, without feedback. They were to press the space bar with the index finger of their dominant hand when seeing a solid line at the beginning of each trial to reveal the following word. Every segment was presented at the centre of the screen. The experiment started with seven practice trials. The reading duration of each segment and the answers to the comprehension questions were recorded. The entire experiment lasted ~ 25 minutes.

2.4. Analysis

Since this study was conducted online, we first evaluated the quality of the data by analysing the accuracy of comprehension questions. The average accuracy for comprehension questions across different sentence conditions was as follows: 87.25% for L1–L1, 85.25% for L1–L2, 80.43% for L2–L1 and 79.74% for L2–L2. We then applied a generalized mixed model using the *lme4* and *lmerTest* packages in the R environment to determine whether sentence types (i.e., L1–L1, L1–L2, L2–L1 and L2–L2) influenced comprehension accuracy and performance over time was also assessed. The model included intercepts for each participant and items as random effects. Compared to L1–L1 sentences, comprehension accuracy was significantly lower for L2–L1 and L2–L2 sentences (L2–L1: Estimate = −0.66, SE = 0.21, $p < .01$; L2–L2: Estimate = −0.70, SE = 0.21, $p < .001$). In addition, there was no significant change in performance over time (Estimate = −0.01, SE = 0.01, $p = 0.25$).

LexTALE performance was scored as the average of the percentage of words correct and the average percentage of non-words correct, which adjusts for the uneven number of words and non-words (see Lemhöfer & Broersma, 2012, for justification of this scoring method). For our sample, the average accuracy was 62.4% (SD = 48.4%) and one participant was excluded from the following analysis of the self-paced reading task due to low accuracy (below 40%). The average accuracy in answering the comprehension questions was 83.2% (SD = 37.4%) and another seven participants with accuracy below 60% were removed. Hence, data from 53 participants were included in the final data analysis. Data points with reading durations outside the range of 100–1500 ms were eliminated. Outliers (reading durations beyond 2.5 standard deviations from the mean) were identified for each participant, condition (L1–L1, L1–L2, L2–L2 and L2–L1) and word position (from WP1 to WP7), and subsequently eliminated. In summary, for WP4, 3,070 trials (83.88% of the total collected data) were included in the analysis reported below, and for WP5, 3,064 trials (83.72%) were included.

Bayesian mixed-linear models (Bürkner, 2017) were constructed for WP4 (switch position) and WP5, with the latter of interest because effects are often delayed in self-paced reading tasks (Jegerski, 2014). A Bayesian mixed-linear model has two main advantages compared to a traditional linear mixed model (Vasishth et al., 2018). First, it enables us to focus on the uncertainty surrounding the estimate of interest. In addition, we can fit 'maximal' models that incorporate full covariances for both by-participant and by-item random effects, which usually generate convergence or singularity problems with a linear mixed model. Thus, Bayesian mixed-linear models allow us to obtain the most conservative results regarding the parameters in light of the data and model at hand. The credible interval (CI) and the probability of direction (*pd*) serve as indices of effect *existence* (i.e., the amount of evidence for the alternative hypothesis). To study the *significance* of an effect (in the sense of 'being worthy of attention' or 'importance' as defined by Makowski et al. (2019a); note that this definition deviates from the meaning of significance in null hypothesis testing), we applied an approach known as the region of practical equivalence (ROPE) analysis (Kruschke, 2014).

[2] The decision to use WP4 was based on the methodology employed by Bultena et al. (2015), which is one of the key references for our study. Selecting WP4 allows for sufficient context both before and after the critical switch position, which is essential for analysing the effects of language switching on sentence comprehension. In addition, the word at the switch position could be any types of word, including verbs, articles, pronouns, adjectives, prepositions and nouns.

https://doi.org/10.1017/S1366728925000100 Published online by Cambridge University Press

The starting point is to establish a 'null effect'. Unlike the frequentist approach that assumes a point null value, in the Bayesian approach, we can define an ROPE for the null region (see more details in Kruschke, 2014 and in Vasishth et al., 2018). The decision on the ROPE range is subjective and based on previous studies but it is not arbitrary. In a self-paced reading task with eye-tracking methodology, Vasishth et al. (2018) suggested a ROPE range of ±20 ms around 0 ms for the total reading time of each segment. Determining the optimal ROPE range lacks a definitive solution, but considering the precedent set by prior pertinent studies, we also applied this range to our results, as the eye-tracking total time and effect sizes in self-paced reading are similar. Moreover, we used the percentage of the full posterior distribution because this might be more sensitive, compared to using either 95% or 89% of CI, and would indicate the portion of the entire posterior distribution in the ROPE (Makowski et al., 2019b).

Due to the positive skewness of the untransformed raw data, reading durations were log-transformed to satisfy the model assumptions. Therefore, the absolute ROPE range of ±20 ms (see above) was also log-transformed in all models. All results presented in the Bayesian tables below are in log format but the descriptive results presented in the main text below were back-transformed to raw reading times for clearer illustration. Four experimental conditions were defined based on the two factors: switch (switch vs. non-switch) and language of the second half of the sentence (Chinese vs. English). In addition, we included L2 proficiency as a variable. All three variables were centred in the models. Models included crossed random intercepts and slopes for switch, language and their interaction, for both participants and items. Analyses were carried out in the *R* environment (R Core Team, 2023) using the packages *brms* (version 2.15.0; Bürkner, 2017) and *bayestestR* (version 0.13.2; Makowski et al., 2019a). The default brms uninformative priors were used.

3. Results

Average reading durations across all conditions and word positions are presented in Figure 1 (panel A); for switch/post-switch positions (WP4 and later), duration differences between switch and non-switch conditions are shown in panel B. Reading durations for WP4 and WP5 under each experimental condition are reported in Table 3. Results from Bayesian mixed-effects models for WP4 and WP5 are presented in Table 4. Each model was executed with six chains of 6,000 iterations. The *Rhat* value for every parameter in each model is 1, indicating successful model convergence. For each analysed model, the median effect estimates, the upper and lower bounds of a 95% CI, the probability that the effect of each variable was >0 or < 0 (probability of direction, *p* [*b* </> 0]) and the percentage of the full CI that is the null region (ROPE) are reported for each variable.

For the interpretation of our results, we followed the Reporting Guidelines for the package bayestestR at https://easystats.github.io/bayestestR/articles/guidelines.html. According to these guidelines, a ROPE value of <1% is considered 'significant', <2.5% is 'probably significant', 2.5%–97.5% suggests 'undecided significance', >97.5% means 'probably negligible' and >99% implies a 'negligible' result. For the probability of direction estimate, a value of ≤95% implies an 'uncertain' effect, >95% suggests 'possibly existing', >97% is 'likely existing', >99% is 'probably existing' and >99.9% is 'certainly existing'. For additional information on ROPE and the probability of direction, please consult Makowski et al. (2019a).

3.1. Word position 4

According to Table 4, the variable switch (switch vs. non-switch) has a probability of 99.59% being positive (Median = 13.6 ms, 95% CI = [3.4 ms, 20.6 ms]), which implies that the effect probably exists. However, the percentage of the full posterior distribution that falls within the ROPE is 95.16%, suggesting that evidence regarding its significance is insufficient. The interaction between language and L2 proficiency has a probability of 99.06% of being negative (Median = −93.8 ms, 95% CI = [−150.5 ms, −19.4 ms]), but is only probably significant (2.19% in ROPE). Regarding this interaction, it is predictable that when reading in English in the latter portion of a sentence, proficiency in L2 (English) will impact reading durations for English; by contrast, when reading in Chinese, L2 proficiency is unlikely to affect reading durations for Chinese. Importantly, the interaction between switch and language has a probability of over 99.99% being negative (Median = −81.5 ms, 95% CI = [−98.5 ms, −60.5 ms]) and is significant (0.00% in ROPE). There is insufficient compelling evidence to support the existence and significance of all other effects; critically, this affects the main effect of L2 proficiency.

Table 3 suggests that language switch incurs a cost when switching into L1, but a benefit when switching into L2 at WP4. We followed up on the interaction between switch and language in Table 4 with a simple effect analysis, which explored the switch effect for each language separately. The results are shown in Table 5. When switching into L1, the effect of the switch has a probability of over 99.99% being positive (Median = 61.8 ms, 95% CI = [46.3 ms, 82.1 ms]) and can be considered significant (0.00% in ROPE). By contrast, when switching into L2, the effect of the switch has a probability of 99.99% of being negative (Median = −31.7 ms, 95% CI = [−46.5 ms, −16.3 ms]), implying that this effect is likely exists but that it is negligible (5.15% in ROPE).

3.2. Word position 5

As shown in Table 4, both the effects of switch (switch vs. non-switch) and of language (Chinese vs. English, language used in the second half of sentence) have relatively high probabilities of being positive (Switch: *pd* = 99.99%, Median = 14.0 ms, 95% CI = [7.0 ms, 24.9 ms]; Language: *pd* = 99.99%, Median = 21.3 ms, 95% CI = [10.5 ms, 32.4 ms]), but they are not significant, with a percentage of 91.85% and 50.56% in ROPE, respectively. The interaction between switch and language has a probability of over 99.99% of being negative (Median = −44.9 ms, 95% CI = [−59.5 ms, −29.6 ms]) and the effect is significant (0.15% in ROPE). There is insufficient evidence to support the existence and significance of all other effects; this regards the main effect of L2 proficiency, as well as the interactions that involve this variable.

Table 3 suggests a cost when switching into L1 but a benefit of a much smaller size when switching into L2 at the WP5. The output of a simple effect analysis, which considered switch effects for each language separately, is shown in Table 5. When switching into L1, the effect of the switch has a probability of over 99.99% being positive (Median = 38.8 ms, 95% CI = [27.8 ms, 50.1 ms]) and the effect can be considered to certainly exist and be significant (0.08% in ROPE). Hence, the switch cost into L1, which emerged for WP4, is also confirmed in WP5. By contrast, when switching into L2, the effect of the switch only has a probability of 90.17% being negative and can be considered uncertain and insignificant (97.14% in ROPE).

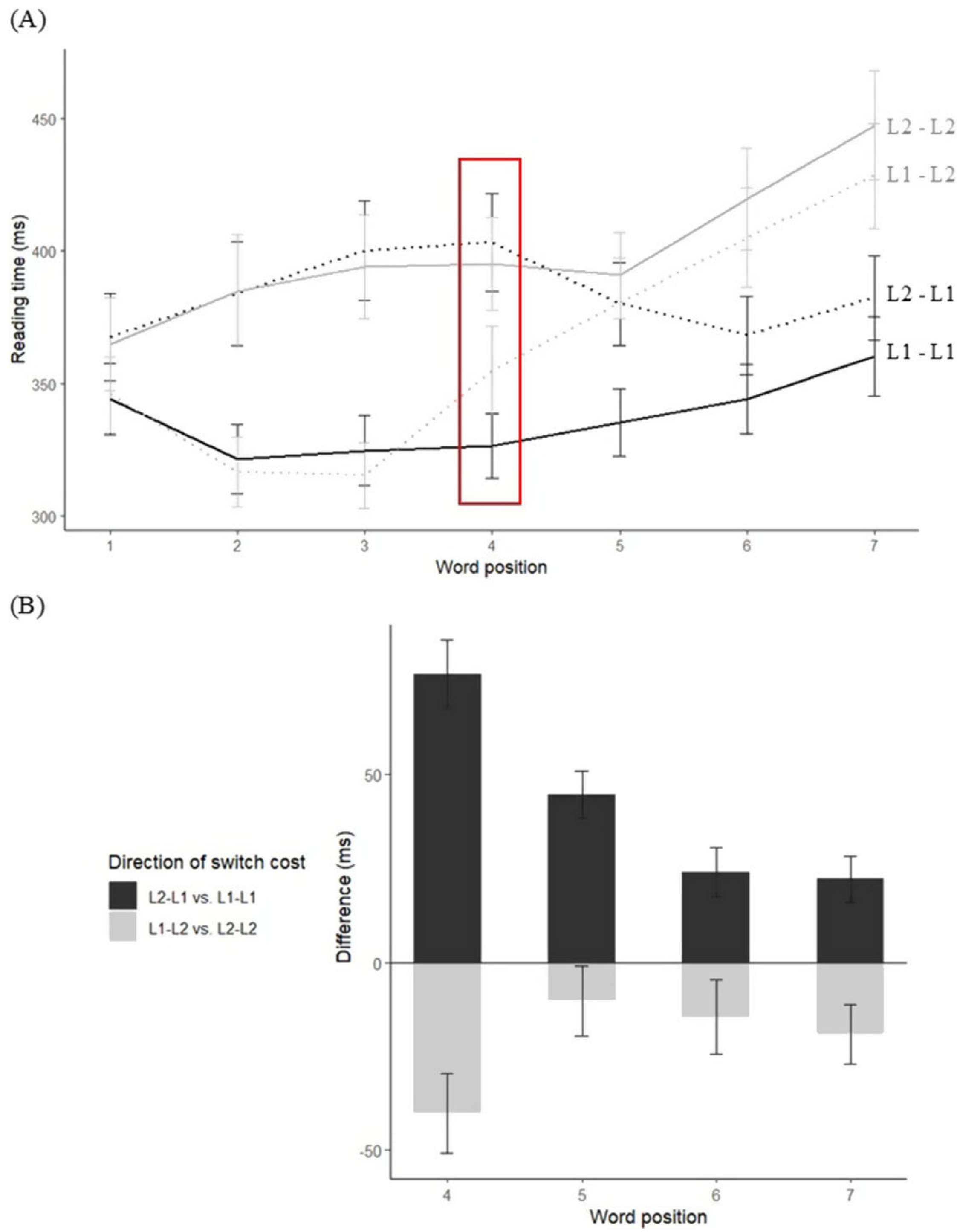

Figure 1. (A) Average reading durations (in milliseconds) for each word dependent on word position within the sentence (1–7) and experimental condition (L2–L2, L1–L2, L2–L1 and L1–L1). The red box indicates the potential language switch position (WP4). Error bars indicate standard errors; reading times are aggregated by participants. (B) Difference between reading durations (switch minus non-switch) dependent on word position (from WP4, switching position, to the end of sentence) and language (switch into L1 vs. into L2), by using the comparison pattern 2.

4. Discussion

The aim of the present study was to elucidate language-switching patterns in comprehension, utilizing cross-script languages in an intrasentential switching scenario. Importantly, employing the self-paced reading paradigm allows us to circumvent the confound of response repetition/switching that previous studies may have encountered (see Introduction). In addition, we maintained consistency in the language of the second half of the sentence across compared patterns, thereby avoiding the confounding influence of a benefit when comparing native to second-language reading. Through this comparative approach, our study elucidated how preceding words in one language within a sentence modulate the

https://doi.org/10.1017/S1366728925000100 Published online by Cambridge University Press

Table 3. Average reading durations (in milliseconds; standard errors in parentheses) for word positions 4 (WP4) and 5 (WP5), dependent on Language (L1; Chinese vs. L2; English) and Switch (non-switch vs. switch)

	L1 (Chinese)		L2 (English)	
	Non-switch L1–L1	Switch L2–L1	Non-switch L2–L2	Switch L1–L2
WP4	326 (4.13)	403 (6.88)	394 (6.66)	356 (5.84)
WP5	335 (4.24)	380 (5.36)	390 (5.92)	381 (5.71)

Note. L1 (Chinese) and L2 (English) denote the language used in the second half of sentences. This language remains consistent across the conditions in comparison pattern 2 in Table 2, which is also the pattern utilized in the current study.

activation of target word candidates from a non-target language. To summarize the results of the present study, we found a pronounced asymmetrical language-switching cost between Chinese and English. Specifically, there was a reliable switching cost when switching into L1, but not when switching into L2. Furthermore, participants' proficiency in L2 was not a driver for language switching in sentence comprehension, implying that the relative lexical activation is not the mechanism underlying the language switch costs in comprehension between Chinese and English.

How do our results relate to previous studies? Language switching within sentences has been explored both in various alphabetic languages (e.g., Bultena et al., 2015; Guzzardo Tamargo et al., 2016;

Table 4. Results from Bayesian mixed-effects models for word positions 4 and 5 (WP4 and WP5), with switch (switch vs. non-switch), language (L1 vs. L2) and L2 proficiency as variables

Word position	Term	Median	95%CI	p (b </> 0)	ROPE (full)
WP4	Intercept	5.81	[5.72, 5.92]	–	–
	Switch	0.04	[0.01, 0.06]	99.59%	95.16%
	Language	0.02	[−0.01, 0.04]	86.36%	99.91%
	L2 proficiency	−0.36	[−1.43, 0.72]	74.44%	7.21%
	Switch × Language	−0.28	[−0.35, −0.20]	> 99.99%	0.00%
	Switch × L2 proficiency	−1.31e−04	[−0.29, 0.30]	50.03%	31.47%
	Language × L2 proficiency	−0.33	[−0.60, −0.06]	99.06%	2.19%
	Switch × Language × L2 proficiency	0.21	[−0.49, 0.92]	72.67%	11.21%
WP5	Intercept	5.84	[5.74, 5.93]	–	–
	Switch	0.04	[0.02, 0.07]	99.99%	91.85%
	Language	0.06	[0.03, 0.09]	99.99%	50.56%
	L2 proficiency	−0.28	[−1.27, 0.74]	70.21%	7.98%
	Switch × Language	−0.14	[−0.19, −0.09]	> 99.99%	0.15%
	Switch × L2 proficiency	0.20	[−0.05, 0.45]	94.44%	11.08%
	Language × L2 proficiency	−0.03	[−0.36, 0.29]	56.96%	28.97%
	Switch × Language × L2 proficiency	0.13	[−0.39, 0.66]	68.55%	15.66%

Note. 'Language' represents the language used in the second half of the sentence. 'Switch', 'Language', and 'L2 proficiency' are all centred in the model. The full posterior distribution is used in the ROPE analysis, and the ROPE range is ±20 milliseconds.

Table 5. Effects of language switching on reading times of word positions 4 and 5 (WP4 and WP5) dependent on language in the second half of the sentence (Chinese vs. English)

Word position	Language	Term	Median	95% CI	p (b </> 0)	ROPE (full)
WP4	Chinese	Intercept	5.81	[5.71, 5.90]	–	–
		Switch	0.17	[0.13, 0.22]	>99.99%	0.00%
	English	Intercept	5.83	[5.72, 5.93]	–	–
		Switch	−0.10	[−0.15, −0.05]	99.99%	5.15%
WP5	Chinese	Intercept	5.81	[5.72, 5.90]	–	–
		Switch	0.11	[0.08, 0.14]	>99.99%	0.08%
	English	Intercept	5.87	[5.77, 5.96]	–	–
		Switch	−0.02	[−0.06, 0.01]	90.17%	97.14%

Note. When the language used in the second half of sentence is Chinese (L1–L1 vs. L2–L1), the language switching direction is switching into L1; and when the language is in English (L2–L2 vs. L1–L2), the language switching direction is switching into L2.

Salig et al., 2023) and with cross-script language combinations (e.g., Hu & Zhao, 2023; Wang, 2015). However, previous studies may have encountered confounding effects (see Introduction and below), leading to inconsistent results. Therefore, we conducted the present study to clarify the effects of language switching on comprehension at the sentence level. Hu and Zhao (2023) utilized a self-paced reading task to investigate language switching between Chinese and English. However, their primary focus was on exploring the role of syntax in language-switching comprehension. Consequently, they manipulated syntax to be consistent or inconsistent with the matrix language. In contrast, given the mixed switching patterns observed from previous studies, our study therefore aimed to clarify the language-switching pattern in comprehension and, to this aim, we employed simple sentence structures that were consistent across Chinese and English.

Bultena et al. (2015) employed a self-paced reading task to investigate language switching between English and Dutch. They observed a significant switching cost only when switching into L2, with L2 proficiency shown to modulate this cost. In contrast, Wang (2015) utilized a maze study to examine language switching between Chinese and English. The author noted distinct switch cost patterns across different word positions: (1) at word position 2 (comparing 'shoe' and '鞋', the same word expressed in two different languages, akin to Comparison pattern 1), a switching cost was evident when switching into L2, modulated by language dominance; and (2) at word position 3 (comparing two 'yesterdays', one positioned after the word in English and the other following the word in Chinese, a parallel to Comparison Pattern 2), both switching directions into L1 and L2 were observed, independent of language dominance. Notably, as mentioned in Introduction, these studies employed different comparison patterns. Hence, the variability in previous results may stem from the differing comparison patterns employed. We will elucidate this aspect first.

4.1. The comparison pattern

As outlined in the Introduction, the comparison pattern utilized could influence the interpretation of language switching in comprehension. The outcome depends on whether the first half of a sentence is held constant and the language in the second half switches or not (as in Bultena et al., 2015, and also in word position two of Wang, 2015; comparison 1 in Table 2), or whether the language in the second half of the sentence is held constant, whereas the language in the first half is potentially different (comparison 2 in Table 2). Bultena et al. (2015) noted this aspect (see their Footnote on p. 462) and provided average reading durations for the position following the switch position. Figure 2 (panel A) compares Bultena's and our results according to comparison pattern 1. What is immediately apparent is that Bultena et al.'s and our results resemble each other strongly: there is a substantial cost when switching from L1 to L2 but much less, or no, effect when switching from L2 to L1. Figure 2 (panel B) replots the same data but now with comparison pattern 2. It is clear that by using this comparison pattern, there is a stronger switching cost when switching to L1, regardless of studies and word positions. When using comparison pattern 1, switching to L2 exhibits a large cost, whereas switching to L1 shows little effect. When using comparison pattern 2, costs of switching to L1 dominate, with switching to L2 either much smaller (as in Bultena et al.) or even numerically reversing to a switch benefit (as in our findings).

To further follow-up this analysis, we conducted an additional statistical analysis in which switching was computed according to comparison pattern 1. The resulting tables are available in the Appendix (from Table 2 to Table 5). Applying the same Bayesian mixed model analysis approach but with an alternative comparison pattern, at word position four (switch position), we observed a significant effect of language (note that in this comparison pattern, language indicates the language used in the first half of sentence) and a significant three-way interaction effect of switch, language and L2 proficiency as elaborated in the Appendix (Table 3). At word position 5 (switch +1), there was a significant interaction effect of switch and language. In summary, when using this comparison pattern, we found a reliable switching cost when switching into L2, with this switching cost being modulated by L2 proficiency. These findings mirror those of previous studies that employed this comparison pattern, specifically, the results reported by Bultena et al. (2015). However, it is crucial to acknowledge the presence of native language benefits which serve as a confounding factor when using this comparison approach. The differential in reading times between L1 and L2 is unsurprisingly affected by participants' proficiency in L2. This confounding variable could be addressed by recruiting balanced bilingual participants in future studies to provide more controlled investigations into language-switching effects.

In addition to mitigating the confounding effect of a native language benefit, adopting comparison pattern 2 aligns with the methodology of previous studies on language switching in production (e.g., Costa & Santesteban, 2004; Johns & Steuck, 2021; Meuter & Allport, 1999; Verhoef et al., 2009). Furthermore, as highlighted by Dijkstra and van Heuven (2002), section 4.4), future studies should explore how the language of preceding words in a sentence can influence the activation of target word candidates from a non-target language. Thus, our study underscores the importance of employing this (in our view) more appropriate comparison pattern in future language-switching research. In Wang's (2015) study, the author applied this comparison strategy at word position 3. However, because in the maze task participants are instructed to select the correct word from two options based on the preceding context, it also involves response repetition/switching between consecutive responses. As outlined in the Introduction, the effects of switching are difficult to interpret in tasks of this type without taking the potential effects of response repetition or switching into account. Our present study addresses this issue by using the self-paced reading task in which response repetition/switching is irrelevant because participants simply press a single key to reveal each upcoming word. In addition, the self-paced reading task allows for the examination of further word positions after the critical switching point.

4.2. The mechanism underlying language switching in comprehension

The primary finding of our study reveals a significant switching cost when switching into L1. Furthermore, L2 proficiency does not appear to modulate switching costs, regardless of switching directions. The concept of inhibition during bilingual language processing has emerged as a prevalent theoretical perspective, with the asymmetrical switch cost being one of the indicators of inhibitory processes (Declerck & Koch, 2023). Therefore, this discussion begins by addressing the potential role of inhibitory control mechanisms.

As outlined in the Introduction, both the ICM and the BIA+ models of bilingual language organization assume that bilinguals co-activate multiple linguistic systems. Both models also agree that

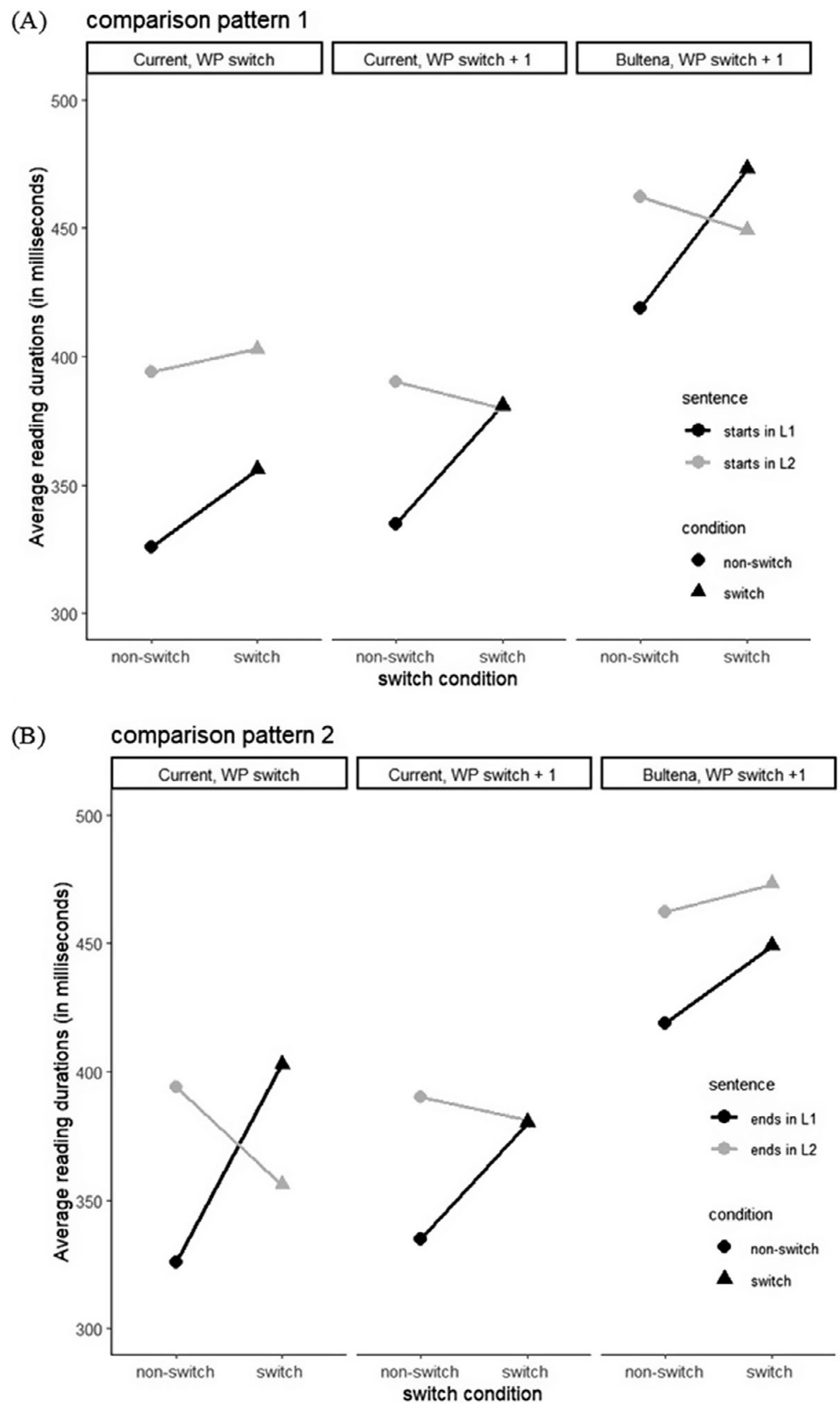

Figure 2. Comparison of results of the present study with those of Bultena et al. (2015) by using comparison pattern 1 (panel A) and comparison pattern 2 (panel B; see Table 2). 'WP switch' indicates the word position at which language switched within target sentences; 'WP switch +1' indicates the following word position

a task or activity involves the formation of task schema (e.g., naming pictures, translating words etc.) and a chosen task schema regulate output from the lexico-semantic system according to the individual's goals in a given context. Maintenance of a task schema involves inhibition of potentially competing task schemas, particularly so for bilinguals: given the assumption that bilinguals co-activate their multiple language systems, maintenance of a task schema (e.g., 'speak in English') involves inhibition of potentially competing task schemas ('speak in French'). Critically, only in the ICM but not in the BIA+ model, inhibition is reactively applied to

https://doi.org/10.1017/S1366728925000100 Published online by Cambridge University Press

the co-activated but non-relevant language system (Declerck & Koch, 2023). This assumption potentially predicts larger switch costs into L1 than into L2: as the L1 lexicon is more accessible than the L2 lexicon, more effort is required for it to be inhibited when processing L2 words than vice versa. When subsequently switching languages, more effort is required to switch back into the (strongly inhibited) L1 than into the (less inhibited) L2. If so, the switch cost asymmetry should be influenced by an individual's L2 proficiency. By contrast, because according to the BIA+ processing within the lexico-semantic system is autonomous and not regulated by the task schema level, switch costs may only arise at the latter level and a potential asymmetry in switch costs is less readily explainable within this framework.

How can these models account for our findings? On the face of it, the absence of an effect of L2 proficiency locates the switch cost outside the linguistic system, because L1/L2 dominance should influence relative mental lexical activation levels and, therefore, lemma retrieval, if inhibition were functioning at that level. Instead, switch costs presumably arise at the task schema level. Interestingly, in our study, participants were explicitly instructed to disregard the presentation language as much as possible when reading each word, and so presumably only a single task schema was formed ('decode the meaning of each stimulus as fast as possible to compute the sentence meaning'). Hence, switch costs in our self-paced reading task cannot easily be accounted for competition between task schemas (however, participants may perceive processing Chinese and English words as distinct tasks, activating competing schemas – 'read in L1' and 'read in L2' – which could explain our results). Furthermore, the task/decision system proposed by the BIA+ model is presumably also affected by non-linguistic factors such as task instructions, demands or participant expectations (Dijkstra & van Heuven, 2002). Non-linguistic context can result in adjustments to decision parameters based on task demands. Perhaps participant expectations may account for the switch costs observed in our study. Specifically, when participants encounter the initial words of a sentence in L2, they may develop an expectation that the sentence will continue in L2. When the sentence unexpectedly switches to L1, this expectation is violated, resulting in temporarily slower reading times compared to sentences that are consistently presented in L1. Conversely, sentences beginning in L1 likely align with participants' 'default' expectations (as words in their native language are processed more naturally), meaning no strong expectation is formed that could be disrupted by a switch to L2. This explanation remains speculative and was formulated post hoc in light of the observed results.

Although the present study found no effect of language proficiency on language switching, this does not fully rule out a linguistic origin of switch costs. The word identification system (in BIA+ model) involves three key factors: lexical, syntactic and semantic sources (e.g., sentence context; Dijkstra & Van Heuven, 2002). The null effect of language proficiency suggests that lexical factors, specifically, the relative activation of words in different languages, are not the primary drivers of switch costs at the identification system level.[3] Another potential linguistic factor to consider is sentence context with regard to syntactic and semantic information; accordingly, the recognition of words within a sentence context is influenced by syntactic and semantic information from both languages (Dijkstra & van Heuven, 2002). As previously noted, a study by Hu and Zhao (2023) investigated the role of syntactic processing in language switching between Chinese and English at the sentence comprehension level. The authors demonstrated that syntactic processing contributes to the costs incurred in code-switching. However, in the present study, all sentence structures were natural and the sentences across language conditions were direct word-for-word translations. Given that the primary aim of this study was to clarify the language-switching pattern in sentence comprehension, rather than to focus on the role of syntax, the current design may not provide the ideal circumstance for thoroughly examining syntactic contributions to switching costs. Nonetheless, syntactic processing may still have played a role, and future research should further explore this question to confirm the influence of syntax in language switching during comprehension, particularly at the sentence level.

A further important point highlighted in the Introduction arises from the central assumption of both the ICM and the BIA+ that bilinguals co-activate their multiple linguistic systems. In orthographic tasks, it is likely that orthographic overlap between the multiple languages is also relevant. Indeed, Dijkstra and van Heuven (2002) explicitly stated that co-activation is less likely in language combinations such as Chinese and English, where the target language is clearly identified by the orthographic script. The influence of language-specific orthographic properties on comprehension during language switching has been previously suggested (e.g., Orfanidou & Sumner, 2005). In their lexical decision task, Greek and English stimuli were constructed from language-specific (unique to one language) and language-nonspecific (common to both languages) letters, and results revealed that language-specific features significantly reduced switch costs without impacting response repetition. Thus, exploring cross-script languages such as Chinese and English offers valuable insights into language switching in comprehension. How would processing in our inter-sentential language-switching task with orthographically dissimilar languages play out in the BIA+? Each incoming word would be processed in the appropriate orthographic stream according to its language, resulting eventually in access to semantics and the exclusive activation of its corresponding 'language node' (i.e., no co-activation). The decision/task schema level would then pick up the activation from the word identification system and use it to fulfil the 'task schema', ultimately resulting in sentence comprehension. If so, the observed switch cost in our experiment could only arise at the decision/task schema level (in line with the finding that L2 proficiency was irrelevant).

Another intriguing finding from our study is a tendency towards facilitation when switching into L2, where the preceding word in L1 aids in processing the target L2 word. As our study does not involve response switches/repeats, this facilitation cannot arise from a confound with response switching (as it may in studies of single word comprehension; see Introduction); however, the benefit when switching into L2 was not robust in our findings. Moreover, as indicated in Table 1 (in the Appendix), there was no interaction between word frequency and the switch factor, suggesting that the facilitation did not arise from the materials used in our study. If this facilitation effect is genuine, it may suggest that when switching into L2, the L2 lexicon is easily activated due to the preceding activation in the L1 lexicon, implying the co-activation of two language nodes. In this scenario, inhibition of the non-target language may not be

[3] To further verify that the observed switch costs are not influenced by the linguistic property of English word frequency, we conducted an additional analysis examining the interaction of word frequency (Zipf frequency sourced from SUBTLEX-UK; van Heuven et al., 2014) and switch (switch vs. non-switch) when the second half of the sentence was in English (L2–L2 and L1–L2). As demonstrated in Table A.1 in the Appendix, no significant effects related to word frequency were observed.

necessary. Given the non-significant nature of the facilitation effects we observed, we propose that this may be due to the absence of expectation interference when the sentence begins in L1.

We acknowledge the limitations of LexTale as a measure of L2 proficiency that focuses solely on lexical knowledge, neglecting other aspects of L2 processing. It also measures accuracy, unlike our study, which examined reaction times – potentially weakening correlations. Future research should consider more comprehensive tools to capture the multidimensional nature of L2 proficiency. More generally, it is recognized that non-significant correlations (in our case, between L2 proficiency and switch cost) make definitive conclusions challenging.

5. Conclusion

In summary, our study elucidates the dynamics of language switching in sentence comprehension. We found a notable asymmetrical switching cost in cross-script language scenarios, specifically, a pronounced cost when switching into L1, but not the reverse. Notably, L2 proficiency did not influence these switching costs, indicating that the *relative lexicon activation level is not the mechanism governing language switching in sentence comprehension between Chinese and English.* Furthermore, our study critically highlights the significance of using different comparison patterns when investigating language switching during sentence comprehension. It will be crucial for future research to adopt consistent and rational comparison patterns, and doing so will yield deeper insights into the mechanisms underlying bilingual language organization more broadly. Future studies should also attempt to provide computational simulations of bilingual language processing which will serve to further validate some of the inferences drawn from empirical studies.

Data availability statement. The data that support the findings of this study are openly available in OSF at https://osf.io/cquk7/?view_only=65c9427416ee4129a2a5da27f235f9af.

References

Adler, R. M., **Valdés Kroff, J. R.**, & **Novick, J. M.** (2020). Does integrating a code-switch during comprehension engage cognitive control? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **46**(4), 741–759. https://doi.org/10.1037/xlm0000755

Altarriba, J., **Kroll, J. F.**, **Sholl, A.**, & **Rayner, K.** (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, **24**(4), 477–492. https://doi.org/10.3758/BF03200936

Alvarez, R. P., **Holcomb, P. J.**, & **Grainger, J.** (2003). Accessing word meaning in two languages: An event-related brain potential study of beginning bilinguals. *Brain and Language*, **87**(2), 290–304. https://doi.org/10.1016/S0093-934X(03)00108-1

Aparicio, X., & **Lavaur, J.-M.** (2014). Recognising words in three languages: Effects of language dominance and language switching. *International Journal of Multilingualism*, **11**(2), 164–181. https://doi.org/10.1080/14790718.2013.783583

Beatty-Martínez, A. L., **Guzzardo Tamargo, R. E.**, & **Dussias, P. E.** (2021). Phasic pupillary responses reveal differential engagement of attentional control in bilingual spoken language processing. *Scientific Reports*, **11**(1), 1–12. https://doi.org/10.1038/s41598-021-03008-1

Bialystok, E., **Craik, F. I. M.**, **Green, D. W.**, & **Gollan, T. H.** (2009). Bilingual minds. *Psychological Science in the Public Interest*, **10**(3), 89–129. https://doi.org/10.1177/1529100610387084

Blanco-Elorrieta, E., & **Pylkkänen, L.** (2018). Ecological validity in bilingualism research and the bilingual advantage. *Trends in Cognitive Sciences*, **22**(12), 1117–1126. https://doi.org/10.1016/j.tics.2018.10.001

Bultena, S., **Dijkstra, T.**, & **Van Hell, J. G.** (2015). Language switch costs in sentence comprehension depend on language dominance: Evidence from self-paced reading. *Bilingualism: Language and Cognition*, **18**(3), 453–469. https://doi.org/10.1017/S1366728914000145

Bürkner, P. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, **80**(1), 1–28. https://doi.org/10.18637/jss.v080.i01

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. https://doi.org/10.1016/j.jml.2004.02.002

Declerck, M., & **Koch, I.** (2023). The concept of inhibition in bilingual control. *Psychological Review*, **130**(4), 953–976. https://doi.org/10.1037/rev0000367

Declerck, M., & **Philipp, A. M.** (2015). A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review*, **22**(6), 1630–1645. https://doi.org/10.3758/s13423-015-0836-1

Declerck, M., **Koch, I.**, **Duñabeitia, J. A.**, **Grainger, J.**, & **Stephan, D. N.** (2019). What absent switch costs and mixing costs during bilingual language comprehension can tell us about language control. *Journal of Experimental Psychology: Human Perception and Performance*, **45**(6), 771-789. https://doi.org/10.1037/xhp0000627

Dijkstra, T., & **van Heuven, W. J. B.**. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, **5**(3), 175–197. https://doi.org/10.1017/S1366728902003012

Drummond, A., **von der Malsburg, T.**, **Erlewine, M. Y.**, & **Vafaie, M.** (2016). Ibex Farm. Retrieved from https://github.com/addrummond/ibex

Forster, K. I., **Guerrera, C.**, & **Elliot, L.** (2009). The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods*, **41**(1), 163 171. https://doi.org/10.3758/BRM.41.1.163

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. https://doi.org/10.5334/joc.186

Goldrick, M., & **Gollan, T. H.** (2023). Inhibitory control of the dominant language: Reversed language dominance is the tip of the iceberg. *Journal of Memory and Language*, **130**, 104410. https://doi.org/10.1016/j.jml.2023.104410

Grainger, J. (1993). Visual word recognition in bilinguals. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon*. John Benjamins Publishing.

Grainger, J., & **Dijkstra, T.** (1992). On the representation and use of language information in bilinguals. In R. J. Harris (Ed.), *Advances in psychology* (Vol. **83**, pp. 207–220). North-Holland. https://doi.org/10.1016/S0166-4115(08)61496-X

Grainger, J., **Midgley, K.**, & **Holcomb, P. J.**, (2010). Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). In M.Hickmann, & K.Michele (Eds.), *Language acquisition across linguistic and cognitive systems* (pp. 267–283). John Benjamins B.V.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, **1**(2), 67–81. https://doi.org/10.1017/S1366728998000133

Green, D. W., & **Abutalebi, J.** (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, **25**(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Guzzardo Tamargo, R. E., **Valdés Kroff, J. R.**, & **Dussias, P. E.** (2016). Examining the relationship between comprehension and production processes in code-switched language. *Journal of Memory and Language*, **89**, 138–161. https://doi.org/10.1016/j.jml.2015.12.002

Hilchey, M. D., & **Klein, R. M.** (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review*, **18**(4), 625–658. https://doi.org/10.3758/s13423-011-0116-7

Hu, W., & **Zhao, Y.** (2023). Code-switching costs from Chinese-English relative clauses processing. *Frontiers in Psychology*, **14**. https://doi.org/10.3389/fpsyg.2023.1144530

Jegerski, J. (2014). Self-paced reading. In J. Jegerski & B. VanPatten (Eds.), *Research methods in second language psycholinguistics* (pp. 20–49). New York: Routledge

Johns, M. A., & **Steuck, J.** (2021). Is codeswitching easy or difficult? Testing processing cost through the prosodic structure of bilingual speech. *Cognition*, **211**, 104634. https://doi.org/10.1016/j.cognition.2021.104634

Koch, I., **Hazeltine, E.**, **Petersen, G.**, & **Weissman, D. H.** (2023). Response-repetition costs in task switching do not index a simple response-switch bias: Evidence from manipulating the number of response alternatives. *Attention, Perception, & Psychophysics*, **85**(8), 2577–2587. https://doi.org/10.3758/s13414-023-02708-2

Koeth, J. T. (2012). *Mix and switch effects in bilingual language processing.* University of Maryland, College Park ProQuest Dissertations Publishing.

Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.

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. https://doi.org/10.3758/s13428-011-0146-0

Liao, W. R., & **Wang, Y.** (2015). The same difference: Comparative syntax-semantics of English same and Chinese Tong/Xiang-Tong. In A. Li, A. Simpson & D. Tsai (Eds.), *Chinese syntax in a cross-linguistic perspective*, pp. 128–147. Oxford: Oxford University Press.

Makowski, D., **Ben-Shachar, M. S.**, **Chen, S. H. A.**, & **Lüdecke, D.** (2019b). Indices of effect existence and significance in the Bayesian framework. *Frontiers in Psychology*, **10**, 14. https://doi.org/10.3389/fpsyg.2019.02767

Makowski, D., **Ben-Shachar, M. S.**, & **Lüdecke, D.** (2019a). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, **4**(40), 1541. https://doi.org/10.21105/joss.01541

Meuter, R. F. I., & **Allport, A.** (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, **40**(1), 25–40. https://doi.org/10.1006/jmla.1998.2602

Ong, G., **McKague, M.**, **Weekes, B.**, & **Sewell, D. K.** (2019). Diffusing the bilingual lexicon: Task-based and lexical components of language switch costs. *Cognitive Psychology*, **114**, 101225. https://doi.org/10.1016/j.cogpsych.2019.101225

Orfanidou, E., & **Sumner, P.** (2005). Language switching and the effects of orthographic specificity and response repetition. *Memory & Cognition*, **33**(2), 355–369. https://doi.org/10.3758/BF03195323

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics*, **18**(7–8), 581–618.

R Core Team (2023). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Rogers, R. D., & **Monsell, S.** (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, **124**(2), 207–231. https://doi.org/10.1037/0096-3445.124.2.207

Salig, L. K., **Valdés Kroff, J. R.**, **Slevc, L. R.**, & **Novick, J. M.** (2023). Linking frequency to bilingual switch costs during real-time sentence comprehension. *Bilingualism: Language and Cognition*, 1–16. https://doi.org/10.1017/S1366728923000366

Struck, J., & **Jiang, N.** (2022). Language switch costs in a lexical decision task: Symmetry and cognitive correlates. *Second Language Research*, **38**(4), 813–838. https://doi.org/10.1177/0267658321998740

Thomas, M. S. C., & **Allport, A.** (2000). Language switching costs in bilingual visual word recognition. *Journal of Memory and Language*, **43**(1), 44–66. https://doi.org/10.1006/jmla.1999.2700

van Heuven, W. J. B., **Dijkstra, T.**, & **Grainger, J.** (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, **39**(3), 458–483. https://doi.org/10.1006/jmla.1998.2584

van Heuven, W. J. B., **Mandera, P.**, **Keuleers, E.**, & **Brysbaert, M.** (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, **67**(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Vasishth, S., **Mertzen, D.**, **Jäger, L. A.**, & **Gelman, A.** (2018). The statistical significance filter leads to overoptimistic expectations of replicability. *Journal of Memory and Language*, **103**, 151–175. https://doi.org/10.1016/j.jml.2018.07.004

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. https://doi.org/10.1016/j.cognition.2008.10.013

von Studnitz, R. E., & **Green, D. W.** (1997). Lexical decision and language switching. *International Journal of Bilingualism*, **1**(1), 3–24. https://doi.org/10.1177/136700699700100102

Wang, X. (2015). Language control in bilingual language comprehension: Evidence from the maze task. *Frontiers in Psychology*, **6**. https://doi.org/10.3389/fpsyg.2015.01179

Wen, Y., & **van Heuven, W. J. B.** (2017). Non-cognate translation priming in masked priming lexical decision experiments: A meta-analysis. *Psychonomic Bulletin & Review*, **24**(3), 879–886. https://doi.org/10.3758/s13423-016-1151-1

Appendix

Table A.1. Interaction between language switch and Zipf frequency of the word used at the corresponding position (WP4 vs. WP5) when English is used in the second half of the sentence (L1–L2, L2–L2)

Word position	Term	Median	95% CI	p (b </> 0)	ROPE(full)
WP4	Intercept	6.03	[5.87, 6.19]	–	–
	Switch	−0.15	[−0.35, 0.04]	93.62%	12.95%
	Zipf frequency	−0.03	[−0.05, −0.01]	99.97%	94.08%
	Switch × Zipf frequency	8.83e−03	[−0.02, 0.04]	70.91%	99.39%
WP5	Intercept	6.02	[5.87, 6.17]	–	–
	Switch	0.06	[−0.10, 0.22]	77.73%	36.44%
	Zipf frequency	−0.02	[−0.04, −0.01]	99.34%	99.39%
	Switch × Zipf frequency	−0.01	[−0.04, 0.01]	86.39%	99.79%

Note. The Bayesian model involved crossed random intercepts and slopes for switch effect, for both participants and items. The full posterior distribution is used in ROPE analysis, and the ROPE range is ±20 milliseconds.

https://doi.org/10.1017/S1366728925000100 Published online by Cambridge University Press

Table A.2. Utilizing comparison pattern 1 (from Table 2), average reading durations (in milliseconds; standard errors in parentheses) for word positions 4 (WP4) and 5 (WP5), contingent upon Language (L1; Chinese vs. L2; English) and Switch (non-switch vs. switch)

	L1 (Chinese)		L2 (English)	
	Non-switch L1–L1	Switch L1–L2	Non-switch L2–L2	Switch L2–L1
WP4	326 (4.13)	356 (5.84)	394 (6.66)	403 (6.88)
WP5	335 (4.24)	381 (5.71)	390 (5.92)	380 (5.36)

Note. The language in the table corresponds to the language utilized in the first half of the sentences. The table presented above aids readers in understanding the functioning of the comparison pattern.

Table A.3. Results from Bayesian mixed-effects models for word positions 4 and 5 (WP4 and WP5) employing comparison pattern 1 from Table 2 which utilizes the language of the first half of the sentence as the “language” factor, with switch (switch vs. non-switch), language (L1 vs. L2) and L2 proficiency as variables

Word position	Term	Median	95%CI	p (b </> 0)	ROPE (full)
WP4	Intercept	5.82	[5.72, 5.92]	–	–
	Switch	0.04	[0.01, 0.06]	99.49%	95.23%
	Language	0.14	[0.10, 0.18]	> 99.99%	0.01%
	L2 proficiency	−0.34	[−1.39, 0.73]	74.32%	7.19%
	Switch × Language	−0.03	[−0.09, 0.03]	85.97%	85.68%
	Switch × L2 proficiency	−1.36e−03	[−0.30, 0.30]	50.37%	31.26%
	Language × L2 proficiency	−0.11	[−0.47, 0.24]	73.37%	22.29%
	Switch × Language × L2 proficiency	0.66	[0.13, 1.21]	99.15%	0.89%
WP5	Intercept	5.84	[5.75, 5.93]	–	–
	Switch	0.04	[0.02, 0.07]	99.99%	91.93%
	Language	0.07	[0.04, 0.09]	> 99.99%	24.80%
	L2 proficiency	−0.25	[−1.26, 0.76]	69.29%	8.62%
	Switch × Language	−0.12	[−0.18, −0.05]	99.98%	3.55%
	Switch × L2 proficiency	0.20	[−0.05, 0.45]	94.28%	10.91%
	Language × L2 proficiency	−0.06	[−0.32, 0.20]	68.48%	31.28%
	Switch × Language × L2 proficiency	0.05	[−0.59, 0.69]	56.53%	14.69%

Note. ‘Switch’, ‘Language’, and ‘L2 proficiency’ are all centred in the model. The full posterior distribution is used in the ROPE analysis, and the ROPE range is ±20 ms.

Table A.4. Bayesian results of the interaction effect between switch and L2 proficiency on word position 4 in different language conditions (the language used in the first half of sentence, comparison pattern 1)

Language	Term	Median	95%CI	p (b </> 0)	ROPE (full)
Chinese	Intercept	5.75	[5.66, 5.84]	–	–
	Switch	0.05	[0.01, 0.09]	99.28%	65.38%
	L2 proficiency	−0.30	[−1.33, 0.72]	71.84%	7.85%
	Switch × L2 proficiency	−0.31	[−0.72, 0.10]	93.26%	7.89%
English	Intercept	5.89	[5.78, 5.99]	–	–
	Switch	0.02	[−0.01, 0.06]	88.86%	97.72%
	L2 proficiency	−0.44	[−1.57, 0.65]	78.49%	6.01%
	Switch × L2 proficiency	0.34	[−0.06, 0.73]	95.16%	5.84%

Note. Since the comparison pattern is based on the language of the first half of sentence, when Chinese is used (L1–L1 vs. L1–L2) the language switching direction indicates switching into L2; and when English is used (L2–L2 vs. L2–L1), the language switching direction is switching into L1.

Table A.5. Effects of language switching on reading times of word position 5 (WP5) dependent on language in the first half of the sentence (Chinese vs. English; comparison pattern 1 in Table 2)

Language	Term	Median	95% CI	p (b </> 0)	ROPE (full)
Chinese	Intercept	5.81	[5.72, 5.90]	–	–
	Switch	0.10	[0.06, 0.14]	> 99.99%	1.70%
English	Intercept	5.87	[5.78, 5.97]	–	–
	Switch	−0.02	[−0.06, 0.02]	80.53%	98.66%

Note. Since the comparison pattern is based on the language of the first half of sentence, when Chinese is used (L1–L1 vs. L1–L2), the language switching direction indicates switching into L2; and when English is used (L2–L2 vs. L2–L1), the language switching direction is switching into L1.