The Modality Switching Costs of Chinese-English Bilinguals in the Processing of L1 and L2

Journal:	*Quarterly Journal of Experimental Psychology*
Manuscript ID	QJE-STD-17-401.R3
Manuscript Type:	Standard Article
Date Submitted by the Author:	30-Aug-2019
Complete List of Authors:	Zhao, Tianyang; South China Normal University, School of Psychology Huang, Yanli; Nanjing Normal University, School of Psychology Chen, Donggui; South China Normal University, School of Psychology Jiao, Lu; South China Normal University, School of Psychology Marmolejo-Ramos, Fernando; University of South Australia, Center for Change and Complexity in Learning Wang, Ruiming; South China Normal University, School of Psychology Xie, Jiushu; Nanjing Normal University, School of Psychology
Keywords:	Embodied Cognition, Conceptual Representation, Switching Costs, Language Production, Bilingual

DOI: 10.1177/1747021819878089

The Modality Switching Costs of Chinese-English Bilinguals in the Processing of L1 and L2

Tianyang Zhao[1†], Yanli Huang[2†], Donggui Chen[1], Lu Jiao[1], Fernando Marmolejo-Ramos[3,4], Ruiming Wang[1*], Jiushu Xie[2*]

[1] Guangdong Provincial Key Laboratory of Mental Health and Cognitive Science, Center for Studies of Psychological Application, School of Psychology, South China Normal University, Guangzhou, China

[2] School of Psychology, Nanjing Normal University, Nanjing, China

[3] Gösta Ekman Laboratory, Department of Psychology, Stockholm University, Stockholm, Sweden

[4] Center for Change and Complexity in Learning, The University of South Australia, Australia

† These authors contributed equally to this work and are co-first authors.

* Correspondence—Ruiming Wang, Guangdong Provincial Key Laboratory of Mental Health and Cognitive Science, Center for Studies of Psychological Application, School of Psychology, South China Normal University, Guangzhou, 510631, Guangdong, China, E-mail: wangrm@scnu.edu.cn, and Jiushu Xie, School of Psychology, Nanjing Normal University, Nanjing, 210097, Jiangsu, China, E-mail: jiusxie@outlook.com.

DOI: 10.1177/1747021819878089

Abstract

Modality switching cost indicates that people's performance becomes worse when they judge sequential information that is related to different sensory modalities than judging information that is related to the same modality. In the present study, we conducted three experiments on proficient and non-proficient bilingual individuals to investigate the modality switching costs in L1 and L2 processing separately. In Experiment 1, materials were L1 and L2 words that were either conceptually related to a visual modality (e.g., light) or related to an auditory modality (e.g., song). The modality switching costs were investigated in a lexical decision task in both L1 and L2. Experiment 2 further explored the modality switching costs while weakening the activation level of the perceptual modality by adding a set of fillers. Experiment 3 used a word-naming task to explore the modality switching effect in language production in L1 and L2. Results of these experiments showed that the modality switching costs appeared in both language comprehension and production in L1 and L2 conditions. The magnitude of the modality switching costs was conditionally modulated by the L2 proficiency level, such as in the L2 condition in Experiment 1 and in both L1 and L2 conditions in Experiment 3. These results suggest that sensorimotor simulation is involved in not only language comprehension but also language production. The sensorimotor simulation that is acquired in L1 can be transferred to L2.

Keywords: Bilingual, Conceptual Representation, Embodied Cognition, Language Production, Switching Costs

DOI: 10.1177/1747021819878089

1 Introduction

How do people represent concepts? Debates over propositional symbol systems and the perceptual symbol systems have been ongoing. It is always a critical theme of conceptual representation in the areas of cognitive science, psychology, education, and neuroscience (Barsalou, 1999, 2008a; Glenberg, Witt, & Metcalfe, 2013; Murphy, 2004; Zwaan, 1999). According to the perceptual symbol systems, people represent concepts via simulating sensorimotor information. Sensorimotor simulations are essential to conceptual processing and representation (Barsalou, 1999). A large body of evidence has shown that the sensorimotor system plays an important role in knowledge acquisition, conceptual processing, and conceptual comprehension, suggesting a close relationship between the perceptual processing and conceptual representation (Ansorge, Kiefer, Khalid, Grassl, & König, 2010; Cai & Connell, 2015; de la Vega, de Filippis, Lachmair, Dudschig, & Kaup, 2012; Marmolejo-Ramos, Khatin-Zadeh, Yazdani-Fazlabadi, Tirado, & Sagi, 2017; Xie, Huang, Wang, & Liu, 2015; Xie, Wang, & Chang, 2014). By contrast, according to propositional symbols systems, concepts are reconstructed into a new representational system. In this system, sensory and perceptual features of concepts are removed. Feature lists, semantic networks, frames, and so on are used to represent the concepts (Fodor, 1975; Pylyshyn, 1984).

Modality switching costs indicate that people have worse performance in judging sequential information that comes from different sensory modalities than judging information from the same modality. Modality switching costs have been widely observed not only in perceptual processing but also in conceptual processing, which

DOI: 10.1177/1747021819878089

might be attributed to the role of sensorimotor simulation in conceptual representation (Pecher et al, 2003).

The conceptual property verification task is often used to test modality switching costs, especially at the conceptual level. In this task, participants verify whether a property (e.g., *rustling*) is a typical feature of a concept (e.g., *leaves*). These properties come from visual, auditory, tactile, gustatory, olfactory, and kinesthetic modalities (see Dagaev & Terushkina, 2014, for the case of sensorimotor modality and affective concepts). Pecher et al.'s (2003) study showed that people responded more quickly and accurately in the property verification task when the property in the previous and current trials (e.g., leaves-rustling) belonged to the same modality (e.g., blender-loud) versus different modalities (e.g., lemon-sour). Pecher, Zeelenberg, and Barsalou (2004) further found the switching costs when participants sequentially judged the same concepts' two properties that came from different modalities. In the study of van Dantzig et al. (2008), participants had to localize either an auditory, visual, or tactile stimulus prior to verify the sensory properties of a concept (e.g., judging whether a banana is yellow). They also found better verifications when the modality of the previously localized stimulus matched the modality of the verified property. Furthermore, Vermeulen, Niedenthal, and Lumineta (2007) found that when participants verified positive and negative words from different sensory modalities, the switching costs have also appeared. These results demonstrated that perceptual information has a facilitation effect on conceptual processing when verifying properties in the same modality than in different modalities. On the contrary, switching among different modalities would produce a processing cost. These results

DOI: 10.1177/1747021819878089

suggested that perceptual simulation might be involved in conceptual representation (see Shen et al., 2016).

However, switching costs not only include facilitation effects that are mentioned above but also include interference effects. For example, in Vermeulen et al's (2008) study, participants were instructed to memorize three (high perceptual sensory load) or one (low perceptual sensory load) sounds (auditory stimuli) or pictures (visual stimuli) before the property verification task (e.g., blender-loud) and then to recall the three or one sound(s) or picture(s) after the verification task. The high (but not low) perceptual load led to slower responses when the yet-to-be-verified properties share the same modality with the memorized perceptual information (i.e., responding whether blender can be loud while memorizing three sounds) than when the yet-to-be-verified properties are in a different modality from the memorized perceptual information (i.e., performing the same task while memorizing three pictures). The findings implied that there were some overlaps between perceptual and conceptual representation, which competed for cognitive resources when they were processed at the same time. The competition effect between perceptual and conceptual processing occurred when the cognitive resources were not sufficient to be used for perceptual and conceptual processing. This was also the reason that the interference effect occurred only in the high perceptual sensory load condition, but not the low (Vermeulen et al, 2008). In other words, the interference effect did not contradict the facilitation effect and both were consistent with the embodied cognition. The interference effect appeared in the high sensory load condition when the competition between the perceptual and conceptual processing occurred, whereas the facilitation effect appeared in the low sensory load condition when the cognitive

DOI: 10.1177/1747021819878089

resources were sufficient for perceptual and conceptual processing. In brief, conceptual processing relies on the perceptual simulation in modality-specific systems.

As mentioned above, the most common paradigm used to test the modality switching cost effects is conceptual properties verification. In this task, participants needed to process the semantic information of words. However, with the semantic activation, there might be a confusion between the perceptual symbol representation and the activated semantic information in the conceptual processing. Therefore, to eliminate the influence of the activation of semantic information, a lexical decision task (i.e., word/non-word judgment) was used, in which participants did not complete semantic processing—that is, the semantic information was only activated at a relatively shallow level (James, 1975).

Increasing evidence shows that language processing relies on perceptual simulation in both L1 (e.g., Barsalou, 1999; Estes, Verges, & Barsalou, 2008; Glenberg & Kaschak, 2002; Hauk, Johnsrude, & Pulvermüller, 2004; Lachmair, Dudschig, de la Vega, De Filippis, & Kaup, 2011; Thornton, Loetscher, Yates, & Nicholls, 2012) and L2 (e.g., Ahlberg, Bischoff, Kaup, Bryant, & Strozyk, 2017; Bergen, Lau, Narayan, Stojanovic, & Wheeler, 2010; De Grauwe, Willems, Rueschemeyer, Lemhöfer, & Schriefers, 2014; Dudschig, de la Vega, & Kaup, 2014; Foroni, 2015; Kühne & Gianelli, 2019). However, Baumeister, Foroni, Conrad, Rumiati, and Winkielman (2017) also pointed out that embodied knowledge that was stored in emotional memory was stronger in L1 than L2. Thus, it is still unclear whether the modality switching costs would occur in L2 processing and whether the L2 proficiency would first have an impact on the sensorimotor simulation in L2 and then on the modality switching costs in L2 (Foroni,

DOI: 10.1177/1747021819878089

2015). Therefore, the purpose of this study is to investigate the modality switching costs of conceptual processing in L2 by adopting a lexical decision task.

Language processing usually consists of language comprehension and language production, which are very different in cognitive mechanisms. Language comprehension is a process of meaning *construction*, with three levels: word comprehension, sentence comprehension, and discourse comprehension. Language production, by contrast, is a process of meaning *expression*, with three stages: conceptualization, language organization, and pronunciation. Although much research has investigated the role of embodiment and perceptual processing in language comprehension, it is seldom investigated in language production.

For language production, two tasks are usually used. One is the word-naming task, in which participants are presented with a word and say the word loudly, whereas, in the picture-naming task, participants are presented with a picture and name the picture loudly (Fraisse, 1967). However, in the picture-naming task, when participants see a picture, they would receive information from all sensory modalities simultaneously. For example, when a participant sees a picture of a train, s/he could see the color of the train and the movement direction of the train. Furthermore, it is difficult to use the picture to present auditory information, such as the train whistle. As we intended to manipulate the modality switching, the picture-naming task is not suitable for the present study. Therefore, we used the word-naming task to explore the sensorimotor simulation in language production in Experiment 3 and to test whether the role of perceptual processing was similar in language comprehension and production.

To recapitulate, the present study used a lexical decision task to investigate

DOI: 10.1177/1747021819878089

whether the modality switching costs could occur in both L1 and L2 lexical comprehension in Experiment 1. Meanwhile, the present study tested whether the modality switching costs would appear when the awareness of perceptual modality information was weakened by adding a set of fillers in Experiment 2. Furthermore, Experiment 3 used a word-naming task to explore the modality switching costs in language production in L1 and L2. According to embodied cognition and perceptual symbol representation, we hypothesized that the modality switching costs would occur in both language comprehension (Experiments 1 and 2) and language production (Experiment 3). According to L2 acquisition theories, the connection between vocabularies and concepts in L1 is more direct and stronger than that in L2. The vocabularies in L2 are connected to concepts with the aid of L1, the strength of which would be increased with the improvement of L2 proficiency (Kroll & Stewart, 1994). For proficient bilinguals, the switching between L1 and L2 would be more fluent than non-proficient bilinguals, which would also contribute to the modality switching of proficient bilinguals. Therefore, we hypothesized that the magnitude of modality switching costs might be modulated by the level of L2 proficiency. That is, the modality switching costs would be weaker for proficient bilinguals than non-proficient bilinguals.

2 Experiment 1

In Experiment 1, Chinese-English bilinguals performed a lexical decision (i.e., word/pseudo-word judgment) task to explore the effects of modality switching in L1 and L2 comprehension.

2.1 Method

2.1.1 Participants

DOI: 10.1177/1747021819878089

Fifty-six college students (age range 19–24, mean age=23, *SD*=1.03, 50 females) took part in Experiment 1. All participants were Chinese-English bilinguals, spoke Chinese (i.e., Mandarin) as the first language and began to learn English in school when they were 12 or 13 years old, with normal or corrected visual acuity. Twenty-eight of them were proficient bilinguals who major in English at a university in China and have passed the Test for English Majors, Grade 4 (TEM-4). TEM-4 is an English proficiency test that English-major college students in China are expected to pass at the end of their second year. They used English in most of their course modules and also in their daily life. The other 28 participants were non-proficient bilinguals who were non-English major university students and none of them have yet passed the College English Test, Band 4 (CET-4). CET-4 is an English proficiency test that non-English major Chinese college students are expected to pass at the end of their second year. They only used English in their English course and seldom used it in their daily life. Participants' selection criteria were the same as those used in Wang, Fan, Liu, and Cai (2014).

2.1.2 Design and Materials

A mixed design of 2 (L2 proficiency: proficient vs. non-proficient, between-subject variable) × 2 (languages: L1 [i.e., Chinese] vs. L2 [i.e., English], within-subject variable) × 2 (modality congruency of word pairs: congruent vs. incongruent, within-subject variable) was used.

Experimental materials are 80 true words, including 40 two-character simplified Chinese words [1] (20 containing visual information and 20 containing auditory

[1] For Chinese characters, there are two versions: traditional Chinese characters, which are usually used in Hong Kong and Taiwan, and simplified Chinese characters, which are

DOI: 10.1177/1747021819878089

information) and 40 English words (20 containing visual information and 20 containing auditory information). Using the 40 Chinese true words, we created ten true word-pairs with the same modality information, including five visual-visual pairs and five auditory-auditory pairs, and ten true word-pairs with different modality information, including five visual-auditory pairs and five auditory-visual pairs. Ten true word-pairs with the same modality information and ten true word-pairs with different modalities were created in the same way for the 40 English words.

In order to ensure that the stimuli were appropriate for the current experiments, we recruited 17 college students who did not take part in the current experiments to do an evaluation task on the degree of visual/auditory information on the visual/auditory-related semantic words on a seven-point Likert scale, with 1 referring to the lowest level of visual/auditory information and 7 referring to the highest level of visual/auditory information. All of the stimuli in the present experiment scored more than 4.5 points. The average score for visual-related words was 5.4 point and the average score for auditory-related words was 5.1, suggesting that all of the visual/auditory words have the dominant information on visual/auditory modality. Meanwhile, the familiarity of materials was scored from the very unfamiliar (i.e., 1 point) to the very familiar (i.e., 7 points). The familiarity of each word was scored more than 5.0 and the mean score was 6.1. Moreover, the average stroke number of Chinese words was 16.8 and the average letter number of English words was 4.9.

Eighty artificial pseudo-words were used as the fillers in the lexical decision task in this experiment. The pseudo-words were composed of two true Chinese characters

---

usually used in mainland China. Our participants are recruited from mainland China and mainly use simplified Chinese characters in their daily life.

DOI: 10.1177/1747021819878089

collocated meaninglessly (e.g., 单简) or an English alphabetic string using English pronunciation rules but without meaning (e.g., latimer). Forty of them were two-character simplified Chinese pseudo-words and the other 40 were English pseudo-words which followed the English pronunciation rule of disyllables or trisyllables. Then, they were created as 20 Chinese pseudo-word-pairs and English pseudo-word-pairs, respectively. Trials with fillers were excluded before analysis.

2.1.3 Procedure

There were two blocks (i.e., a Chinese block and an English block) in this experiment. The presentation order of the two blocks was counterbalanced between the participants. In each block, the 20 experimental word-pairs and the 20 filling word-pairs were mixed and presented in random order.

The experimental procedure and stimuli presentation were controlled by computers using E-prime 1.2 (Psychology Software Tools Inc., Pittsburgh, USA). Participants were tested individually in a dimly lit room. Each trial began with a black cross as the fixation at the center of the screen for 500 ms, followed by the first word of the word-pair at the center of the screen for 3000 ms or until participants' responses. Participants were instructed to respond as quickly and accurately as possible by pressing the F or J key on the keyboard to indicate whether the word was a real word or not. After that, a blank screen appeared for 500 ms followed by a 500-ms black cross, then the second word of the word-pair appeared at the center of the screen for 3000 ms or until participants' responses. Participants were asked to respond as quickly and accurately as possible by pressing the F or J key on the keyboard to indicate whether the word was a real word or not. The allocation of true words or pseudo-words to F or J was

DOI: 10.1177/1747021819878089

counterbalanced between participants. Then, a 500 ms blank screen was presented as the intertrial interval. Response times (RT) and accuracy data were recorded automatically. A practice containing 6 real words and 6 pseudowords was given to help participants familiarize with the procedure. These 12 words were different from those used in the experiment session (see Figure 1).

[Insert Figure 1 here]

2.2 Results

As with previous studies, all participants' accuracy was higher than 75%, and no participants were removed (Connell, 2007; Gozli, Chasteen, & Jay, 2012; Wang, Chi, Wang, & Wu, 2005). Only RTs on the second-word stimuli in every word pair in which participants responded accurately were retained for analysis. For RT analyses, incorrect trials were excluded and correct trials beyond ±3 *SD*s were considered as outliers and were removed, which accounted for 1.88% of the total data (see Marmolejo-Ramos, Cousineau, Benites, & Maehara, 2015). All accuracy data were included in the accuracy analyses. Detailed descriptive results are shown in Table 1.

[Insert Table 1 here]

Analyses were conducted using mixed-effects models with crossed random effects for subjects and items using the lme4 package (Bates, Sarkar, Bates, & Matrix, 2007) and the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2014) in the

DOI: 10.1177/1747021819878089

statistical software R (version 3.4.3). Mixed-effects models are preferable to ANOVA because they allow random effects of participants and items to be considered simultaneously, making the data modeling more appropriate and the results generalizable to studies with similar subjects and items. As recent studies found that using backward model comparison in mixed-effects models might result in non-convergences, we used forward model comparison to determine the best fitting random effect structure in the present study (Bates, Mächler, Bolker, & Walker, 2015; Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017).

In the RT analyses, we employed a mixed effect model with L2 proficiency (proficient vs. non-proficient), word language (l1 vs. l2), modality congruency (congruent vs. incongruent), and their interactions as fixed effects. Regarding random effects, we included by-participant and by-item random intercepts; by-participant random slopes for the main effect of modality congruency and the interaction between modality congruency and word language; and by-item random slopes for L2 proficiency. The other factors and the interaction among the three factors were excluded from the best-fitted model because they did not improve the model fit (ps > .05) (see Hsu & Novick, 2016). For this model, all three variables were coded using mean-centered contrast coding in order to yield results analogous to those obtained from ANOVA.

Table 2 summarized the model for RT. As shown in Table 1 and 2, the effect of L2 proficiency was significant ($t = 2.66$, $p = .01$), suggesting the RT of proficient bilinguals (629 ms) was shorter than that of non-proficient bilinguals (689 ms). The effect of modality congruency reached significance level ($t = -2.63$, $p = .01$), indicating that participants responded faster in congruent word trials (639 ms) than in incongruent word

DOI: 10.1177/1747021819878089

trials (679 ms). The effect of word language did not reach significance (t = -.69, p = .49). Considering our theoretical interest in modality switching costs (i.e., the difference between modality congruent trials and modality incongruent trials), we mainly focused on the interaction results related to the variable of modality congruency. The two-way interaction between L2 proficiency and modality congruency did not reach significance (t = -.28, p = .78).

Critically, the three-way significant interaction between L2 proficiency, modality congruency, and word language was significant (t =2.27, p = .03). Follow-up analyses first showed that the modality congruency × word language two-way interaction was significant (t = -2.12, p = .04) for proficient bilinguals, but not for non-proficient bilinguals (t = .30, p = .76). Specifically, for proficient bilinguals, the modality switching costs were significantly smaller in the L1 (11 ms) than L2 (74 ms) condition. Second, the L2 proficiency × modality congruency two-way interaction was significant in the L1 word condition (t = -2.14, p = .03), but not in the L2 word condition (t = -1.19, p = .25). Specifically, in the L1 word condition, the modality switching costs were significantly smaller for proficient bilinguals than non-proficient bilinguals (See Table 2 and Figure 2).

[Insert Table 2 and Figure 2 here]

Accuracy data were analyzed via a generalized linear mixed effects model with a binomial distribution, with L2 proficiency (proficient vs. non-proficient), word language (L1 vs. L2), modality congruency (congruent vs. incongruent), and their interactions as

DOI: 10.1177/1747021819878089

fixed effects. As random effects, we included by-participant and by-item random intercepts. The other factors and the interaction among the three factors were excluded in the fitted model because they did not improve the model fit (*p*s > .05) (see Hsu & Novick, 2016). For this model, all three variables were coded using a mean-centered contrast.

Table 2 summarized the model for accuracy data. As shown in Table 1 and 2, the effect of L2 proficiency was significant (z = -2.57, p = .01), suggesting that the accuracy of proficient bilinguals (96.96%) was greater than that of non-proficient bilinguals (92.68%). The interaction between L2 proficiency and word language was significant (z = -2.96, $p < .01$). Further analyses showed that, for proficient bilinguals, the accuracy difference between L1 word conditions (97.32%) and L2 word conditions (96.61%) was not significant (p = .84); whereas for non-proficient bilinguals, the accuracy in L1 word condition (96.96%) was significantly larger than the accuracy in L2 word conditions (88.39%) ($p < .01$). Regarding the modality switching costs we focused on, the effect of modality congruency was not significant, nor were the L2 proficiency × modality congruency, modality congruency × word language, and L2 proficiency × modality congruency × word language interaction effects in the accuracy analysis. These results showed that there was no trade-off phenomenon between accuracy and RT so that the effect of modality switching in RT was not at the cost of reducing accuracy.

In brief, for non-proficient bilinguals, the significant main effect of modality congruency showed that the modality switching costs existed in both L1 and L2 word conditions. For proficient bilinguals, the magnitude of modality switching costs was smaller in L1 than L2 word conditions showing the modulation effect of word language in modality switching costs among proficient bilinguals. Similarly, in the L2 word

DOI: 10.1177/1747021819878089

condition, the significant main effect of modality congruency showed that participants responded faster and more accurately to congruent word pairs than incongruent word pairs, suggesting the modality switching costs existed for both proficient and non-proficient bilinguals. In the L1 word condition, the magnitude of modality switching costs was smaller for proficient bilinguals than for non-proficient bilinguals, suggesting that the magnitude of the modality switching costs could be modulated by the level of L2 language proficiency in L1 word conditions.

However, it should be noted that all words used in Experiment 1 were the visual-related or auditory-related semantic words which contained obvious perception modality information. Thus, one may argue that these findings were caused by the awareness of perception modality information, which activated the perception modality strategically during conceptual processing. Therefore, in Experiment 2, we further explored the modality switching effect by adding true words without obvious perceptual information as filling materials to weaken the strategic activation of semantics.

3 Experiment 2

3.1 Method

3.1.1 Participants

Sixty-four college students (age range 19–30, mean age=22, *SD*=2.14, 50 females) took part in Experiment 2, none of whom had participated in Experiment 1. All participants were Chinese-English bilinguals, spoke Chinese (i.e., Mandarin) as a first language and began to learn English in school when they were 12 or 13 years old, with normal or corrected visual acuity. Thirty-two of them were proficient Chinese-English bilinguals and the other 32 were non-proficient Chinese-English bilinguals. The

DOI: 10.1177/1747021819878089

recruitment criteria for proficient and non-proficient bilinguals were the same as those in Experiment 1.

3.1.2 Design and Materials

The design of Experiment 2 was the same as Experiment 1. In Experiment 2, we used the same 80 true words as experimental materials and 80 pseudo-words as filling materials with Experiment 1. In addition, 40 true words without perceptual information were added as filling materials, including 20 two-character simplified Chinese words and 20 English words, which were combined as 10 Chinese word pairs and 10 English word pairs, respectively. The same evaluation task on the degree of visual/auditory information was conducted. All 40 filling words were scored below point 3, which made sure the 40 filling words did not contain typically visual or auditory information. Trials with fillers were excluded before analysis.

3.1.3 Procedure

The procedure was same as Experiment 1 except for the addition of 10 filling word-pairs in both the Chinese block and the English block, resulting in 20 experimental word pairs, 20 pseudo filling word pairs and 10 true filling word pairs in Experiment 2, which were presented in a random order in the Chinese block and English block respectively. The presentation order of the block was counterbalanced between participants.

3.2 Results

The same trimming procedure with Experiment 1 was adopted in Experiment 2. All participants' accuracy was higher than 75% and none of their data was removed. Only responses to the second-word stimuli in each word pair were included for RT and

DOI: 10.1177/1747021819878089

accuracy analyses. For RT analyses, incorrect trials were excluded and correct trials beyond ±3 *SD*s were considered as outliers and removed (4.82% of all responses). All accuracy data were included in the accuracy analyses. Detailed descriptive results are shown in Table 3.

[Insert Table 3 and Figure 3 here]

We employed a mixed-effect model with L2 proficiency (proficient vs. non-proficient), word language (L1 vs. L2), modality congruency (congruent vs. incongruent), and their interactions as fixed effects. As random effects, we included by-participant and by-item random intercepts, with by-participant random slopes for word language. For this model, all three variables were coded using mean-centered contrast coding.

Table 4 summarized the model for RT. As shown in Tables 3 and 4, the effect of L2 proficiency was significant (t = 3.55, p < .001), suggesting that proficient bilinguals (568 ms) responded faster than non-proficient bilinguals (631 ms). The effect of modality congruency was significant (t = -3.28, p = .001), indicating that participants responded faster in congruent word trials (591 ms) than in incongruent word trials (608 ms). The effect of word language was significant (t = 2.31, p = .03), showing that participants responded faster in L1 word conditions (583 ms) than in the L2 word conditions (626 ms). The interaction between L2 proficiency, modality congruency, and word language did not reach significance (t = .99, p = .32).

[Insert Table 4 here]

DOI: 10.1177/1747021819878089

For accuracy data, a generalized linear mixed effects model with a binomial distribution was conducted with L2 proficiency (proficient vs. non-proficient), word language (L1 vs. L2), modality congruency (congruent vs. incongruent), their interactions as fixed effects, and by-participant intercepts, by-item random intercepts, by-participant random slopes for word language as random effects. For this model, all three variables were coded using mean-centered contrast coding. Table 4 summarizes the model for accuracy data. Results showed that the effect of word language was significant ($z = -2.48$, $p = .01$), suggesting the accuracy with Chinese words (M = 98.20%) was more accurate than with English words (M = 95.32%). Accuracy analyses showed that there was no trade-off phenomenon between accuracy and RT so that the effect of modality switching in RT was not at the cost of reducing accuracy.

In Experiment 2, the modality switching costs in lexical decision task was also observed even when the strategic activation of perceptual information was weakened in both L1 and L2 word conditions for proficient and non-proficient bilinguals. Unlike in Experiment 1, the three-way interaction was not significant in Experiment 2. As the only difference between Experiments 1 and 2 is whether experimental materials include fillers, the absence of the 3-way in Experiment 2 may be because the sensorimotor information simulation is weakened by adding fillers.

Language processing includes two different cognitive processes: language comprehension and language production. However, as mentioned in the introduction, most of the previous studies focused on language comprehension to test the underlying process of language processing; how perceptual processing works in language production

DOI: 10.1177/1747021819878089

was rarely discussed. Therefore, Experiment 3 investigated whether the language production process would be affected by perceptual processing and a word-naming task was employed to explore modality switching costs in language production.

4 Experiment 3

4.1 Method

4.1.1 Participants

Fifty-six college students (age range 19–25, mean age=22, SD=1.35, 43 of them were female) took part in Experiment 3, none of whom had participated in Experiments 1 and 2. All participants were Chinese-English bilinguals, spoke Chinese (i.e., Mandarin) as a first language and began to learn English in school when they were 12 or 13 years old, with normal or corrected visual acuity. Twenty-eight of them are proficient Chinese-English bilinguals and the other 28 of them are non-proficient Chinese-English bilinguals. The recruitment criterions for proficient and non-proficient bilinguals are the same as those in Experiment 1.

4.1.2 Design, Materials, and Procedure

The design and stimuli in Experiment 3 were exactly the same as Experiment 2 except that participants were instructed to do a word-naming task, instead of a lexical decision task, and the pseudo filling word pairs were excluded.

The procedure was similar to Experiment 2, which included a Chinese block and an English block. In each block, the 20 experimental true word pairs and 10 filling true word pairs were mixed and presented randomly in each block. The presentation order of two blocks was counterbalanced between the participants.

The stimuli presentation and data collection for RT were controlled by computers

DOI: 10.1177/1747021819878089

with E-prime1.2, while the computer was connected to the SR-Box with a microphone to detect participants' oral responses. Accuracy was recorded by the experimenter manually by counting the number of correct and incorrect trials. Participants were tested individually in a dimly lit room. The procedure for each trial was the same as that in Experiments 1 and 2, except that participants were instructed to name the presented word as quickly and accurately as possible in the corresponding language by the microphone.

4.2 Results

The same trimming procedure was used in Experiment 3. All participants' accuracies were higher than 75% and none of their data was removed. Only responses to the second-word stimuli in each word pair were included for RT and accuracy analyses. For RT analyses, incorrect trials were excluded and correct trials beyond ±3 *SD*s were considered as outliers and removed (4.05% of all responses). All accuracy data were included in the accuracy analyses. Detailed descriptive results are shown in Table 5.

[Insert Table 5 here]

We employed a mixed-effect model with L2 proficiency (proficient vs. non-proficient), word language (L1 vs. L2), modality congruency (congruent vs. incongruent), and their interactions as fixed effects. As random effects, we included by-participant and by-item random intercepts, by-participant random slopes for word language and the interaction between modality congruency and word language, and by-item random slopes for L2 proficiency. For this model, all three variables were coded using mean-centered contrast coding.

DOI: 10.1177/1747021819878089

The model parameters were shown in Table 6, which showed that the effect of L2 proficiency was significant (t = 4.00, p < .001), suggesting proficient bilinguals (632 ms) responded faster than non-proficient bilinguals (712 ms). The effect of modality congruency was significant (t = -3.70, p < .001), indicating that participants responded faster in congruent word trials (665 ms) than in incongruent word trials (679 ms). The effect of word language was significant (t = 4.11, p < .001), indicating that participants responded faster to Chinese words (611 ms) than to English words (733 ms).

The interaction between L2 proficiency and word language was significant (t = 3.94, p < .001). Further analyses first showed that the RT difference (38 ms) between proficient bilinguals and non-proficient bilinguals was smaller in L1 word condition than the difference (123 ms) in L2 word condition. Second, the RT difference (80 ms) between L1 word condition and L2 word condition was smaller for proficient bilinguals than the difference for non-proficient bilinguals.

The two-way interaction (L2 proficiency × modality) was significant (t = -2.18, p = .03). Further analyses showed that the difference of the modality switching costs (12.70 ms) was smaller for proficient bilinguals than for non-proficient bilinguals (16 ms). The three-way (L2 proficiency × modality congruency× word language) interaction was not significant (t = -1.42, p = .16).

For the accuracy data, as it was recorded by the experimenter manually through counting the number of correct and incorrect trials, it was not suitable for further analyses using the generalized linear mixed-effects model. Thus, we only reported descriptive data for accuracy. From the descriptive data, we could see that there was no trade-off phenomenon between accuracy and RT data.

DOI: 10.1177/1747021819878089

[Insert Table 6 here]

In brief, in Experiment 3, the significant main effect of modality congruency in RT analysis suggested that we also found the modality switching costs in language production when the word-naming task was used. The significant two-way (L2 proficiency × modality) interaction further suggested that the modality switching costs in the word-naming task were modulated by the L2 proficiency level. The magnitude of the modality switching costs was smaller for proficient bilinguals than non-proficient bilinguals.

5 Discussion

In the present study, we conducted three experiments among Chinese-English bilinguals with proficiency or non-proficiency level of L2 (i.e., English) to investigate the modality switching costs in L1 and L2 processing separately. In Experiment 1, we found that the modality switching costs existed in language processing with low semantic activation in the lexical decision (word/pseudo-word judgment) task. The modality switching costs existed in both L1 and L2 processing. The observed modality switching costs for non-proficient bilinguals in L2 processing suggested that the sensorimotor simulation in L1 can be transferred to L2. In addition, the modality switching costs in L1 word conditions were smaller for proficient bilinguals than for non-proficient bilinguals, showing the modulation role of the L2 proficiency level in the modality switching costs in the L1 word condition. In Experiment 2, the modality switching costs remained when including a set of true words without obvious perceptual information that weakened the

DOI: 10.1177/1747021819878089

strategic activation of semantics. The L2 proficiency level no longer modulated the magnitude of the modality cost effects. The modality switching costs existed in both L1 and L2 words conditions among proficient and non-proficient bilinguals. The findings in Experiments 1 and 2 combined to show that the modality switching costs existed in language comprehension even with low semantic activation in both L1 and L2. The modulation effect of L2 proficiency disappeared when perceptual information was weakened by adding a set of words without perceptual information. In Experiment 3, the modality switching cost effects were also observed in language production in a word-naming task. The magnitude of the modality switching costs was smaller for proficient bilinguals than for non-proficient bilinguals in both L1 and L2 word conditions. In sum, the modality switching costs could occur not only in language comprehension but also in language production, suggesting that the sensorimotor simulation was involved in language processing and that conceptual and perceptual information may be mutually dependent in language processing.

Besides, the present study used lexical decision and word naming tasks to test switching costs, which can test switching costs at different semantic levels. Previous studies usually used property verification task to test switching costs. This task requires participants to judge whether one property belongs to one concept (Pecher, Zeelenberg, & Barsalou, 2003,2004). Thus, property processing is demanded in previous studies and we do not know whether switching costs would occur when the property is less demanded, such as the lexical decision task. Furthermore, the present study also tested switching cost in L2. While previous studies on sensory symbol systems mainly focus on L1, fewer studies have tested it in L2 (van Dantzig et al., 2008; Vermeulen et al., 2007). The

DOI: 10.1177/1747021819878089

present study cannot only contribute to the understanding of switching costs but also to the understanding of the conceptual representation of bilinguals.

5.1 Modality Switching Cost in Language Comprehension

Regarding the paradigm to test the modality switching costs in language comprehension, although the conceptual-properties verification task was one of the most commonly used tasks, participants have to process the semantic information of the words in such tasks, which led to a confusion between the perceptual symbol representation and the activated semantic information in the conceptual processing. Therefore, the present study adopted a lexical decision task (i.e., word/pseudo-word judgment) in Experiments 1 and 2 in which participants did not need to do semantic processing. Results of these experiments also found switching costs when the semantic processing was not demanded. This may suggest that the modality switching effect was mainly attributed to the activated perceptual information instead of to semantic information. As introduced, all words used in Experiment 1 were visual- or auditory-related semantic words. The obvious perceptual modality information might have made participants strategically activate the perceptual and sensorimotor information when processing the concepts. To avoid this possibility, Experiment 2 added a set of filler stimuli without obvious perceptual information. The findings in Experiment 2 further confirmed that modality switching costs are attributable to the perceptual symbol representation.

Consistent with the previous studies, perceptual processing was involved in language processing and the representation of concepts was involved in the sensorimotor system and perceptual symbols (Barsalou, 2008b; Casasanto, 2009; Vallesi, Mcintosh, & Stuss, 2011). When information is processed in one specific perceptual modality, this

DOI: 10.1177/1747021819878089

perceptual modality would be activated, which facilitated the processing of the information that comes from the same modality. On the contrary, the activated modality would impair the processing of the information that comes from a different modality. The present study found that, in language processing, when participants processed a word containing visual information, the processing of the following word containing auditory information was impaired, and vice versa (i.e., switching costs). This is also consistent with the embodied cognition, which holds that the sensorimotor simulation is involved in language processing (see Stins, Marmolejo-Ramos, Wenker, Hulzinga, & Cañal-Bruland, 2017). The modality switching costs that were found in perceptual stimuli (e.g., see a picture and then hear a sound) in previous studies were also observed in language processing (e.g., see a word containing visual information and then see a word containing auditory information) in the present study (Barsalou, 2009; Pecher, Zeelenberg, & Barsalou, 2003; Spence, Nicholls, & Driver, 2001).

5.2 Modality Switching Cost in Language Production

The present study also investigates whether the modality switching cost effect would occur in language production in Experiment 3. Specifically, we used the word-naming task, in which participants watched a word and asked to say the word directly, to explore the sensorimotor simulation in language production.

Combining the results of the three experiments in the present study, we found that the modality switching costs could occur in both language comprehension and language production. The findings further confirmed the views of perceptual symbol theory that sensorimotor simulations were involved in the conceptual representations.

As stated, there are two basic levels for language representation: conceptual level

DOI: 10.1177/1747021819878089

and lexical level (Li, Mo, & Wang, 2009; Zeelenberg & Pecher, 2003). The process of language production usually involved a process from the conceptual level to the lexical level (Silverberg & Samuel, 2004). The lexical generation was considered as the core component of language production. Specifically, in the lexical level, people first select the suitable lemma (i.e., the semantic-syntactic representation of a word), then add the sound representation (i.e., the phonological encoding) to the lemma, and finally produce the word through articulation with the necessary muscles (Declerck & Philipp, 2015; Levelt et al., 1999). The observed modality switching costs in language production in Experiment 3 suggested that the sensorimotor simulations were also spontaneously involved in language production. Further studies may test in which stage(s) the sensorimotor simulations were involved.

5.3 Comparison of the Modality Switching Cost in L1 and L2 processing

Previous studies have a controversy about the simulation of the motion-perception process in L1 and L2. On the one hand, some researchers indicated that the simulation of the motion-perception process in L1 and L2 was similar. Dudschig et al. (2014) compared the activation of spatial information caused by nouns in L1 and L2. In their experiment, German-English bilinguals were asked to respond to nouns using upward or downward arm movements. These nouns involved spatial information and were presented in L1 and L2. Although the meaning in the lexical level was an irrelevant variable, they found activation of action occurred spontaneously in L2, similar to L1, indicating the simulation of motion-perception in both L1 and L2. Grauwe et al (2014) used Dutch and German as experimental materials and showed that there were similar brain activation patterns in high proficiency bilinguals and monolinguals. Both responses of L1 and L2

DOI: 10.1177/1747021819878089

activated perceptual-motor areas, supporting the finding that there was no difference between the representation of L1 and L2.

On the other hand, some researchers pointed out that the process of comprehension in L2 was relatively shallow compared to L1 (Clahsen & Felser, 2006). L1 comprehension could activate the perceptual-motor information because of the strong association of concept and perception, while the L2 comprehension depended more on the word association between L2 lexical form and L1 translation equivalent. Vukovic and Shtyrov (2014) also found the difference between the simulation of the motion-perception process in L1 and L2.

In the present study, Chinese-English bilinguals were asked to respond to the materials in Chinese and English to explore the modality switching costs in L1 and L2. Results showed that the modality switching costs existed in both L1 and L2. In Experiment 1, the modality switching costs were observed in L2, which were similar for proficient and non-proficient bilinguals, showing that the sensorimotor simulation in L2 could occur spontaneously even it needs to be transferred from L1. Similarly, the modality switching costs were also observed in L2 in the word-naming task in language production. These were also in line with the bilingual representation model. In other words, the conceptual representation in L1 and L2 shared many specific concepts (Kroll & Stewart, 1994), but they might be separated at the level of abstract concepts and some concepts with cultural features (Degroot, Dannenburg, & Vanhell, 1994; Dong, Gui, & Macwhinney, 2005). Indeed, as recently shown by Villani, Lugli, Liuzza, and Borghi (in press), situating concepts within a concretness-abstractness dimension is too simplistic; instead concepts, particularly abstract, can be classified in other dimensions such as

DOI: 10.1177/1747021819878089

degree of "sensorimotorness" and degree of inner and social experiencing. In brief, the findings may suggest that the representation of L1 and L2 is partially overlapped which leads to the difference of sensorimotor simulation in L1 and L2 processing.

5.4 The Modulation Role of L2 Proficiency and Word Language

In Experiment 1, we also found that L2 proficient and word language played modulation roles in modality switching costs. Specifically, in the L2 word condition, the modality switching costs were significantly larger for proficient bilinguals than non-proficient bilinguals, whereas the modality switching costs were significantly smaller for proficient bilinguals than non-proficient bilinguals in the L1 word condition. The larger modality switching costs for proficient bilinguals than non-proficient bilinguals was consistent with the L2 acquisition theories that the connections between vocabularies with concepts in L2 was more direct and stronger for proficient bilinguals than for non-proficient bilinguals.

By adding a set of true words without obvious perceptual information as filling materials to weaken the strategic activation of semantics in Experiments 2, the modulation effect of L2 proficiency disappeared. The mentioned strategic activation in Experiment 1 might, at least partially, impair the modality switching costs among non-proficient bilinguals. According to Kroll, Van Hell, Tokowicz, and Green (2010)'s revised hierarchical model, the L1 translation equivalent was required as mediation to access L2 meaning for non-proficient bilinguals, whereas the L2 meaning could be accessed directly via the shared conceptual representations with L2 among the proficient bilinguals. Therefore, it was assumed that the strategies might have an impact on the L2 language processing, especially for non-proficient bilinguals. However, it was not clear

DOI: 10.1177/1747021819878089

how the language processing and representation for non-proficient bilinguals were influenced by such strategies when all words had perceptual modality information. It would be worthwhile to further investigate this in future studies.

5.5 The Modulation Role of L2 Proficiency in Language Production

Experiment 3 tested the modulation effect of the L2 proficiency on the switching cost effect in language production. Results showed this modulation effect even when words without obvious perceptual information were included as fillers, which was different from the results of language comprehension in Experiment 2. This suggests that the modulation effect of the L2 proficiency less relied on the strategic activation and the salience of perception modality in language production is larger than that in language comprehension. In other words, this may suggest that the activation level of the perceptual modality information was stronger in language production than language comprehension. From the view of integrated theory, language production involves comprehension processes and is a form of action (Pickering & Garrod, 2013). Thus, language production would recruit more sensorimotor simulation compared with language comprehension.

Furthermore, this modulation effect of the L2 proficiency was not shaped by testing language (i.e., L1 or L2). This may suggest that, in language production, sensorimotor simulating occurs in the conceptualization stage because this stage does not rely on the lexicon. Pickering and Garrod (2013) also pointed out that language process could be treated as a process that maps from a “higher” (e.g., syntax) to a “lower” (e.g., phonology) linguistic level. Therefore, conceptualization is vital for language production. While perceptual symbol systems hold that conceptual processing is based on

DOI: 10.1177/1747021819878089

sensorimotor simulation, the conceptualization stage is most probably embodied. However, these are post-hoc explanations and future studies may investigate how the modality switching costs occur in language production. Specifically, whether the switching costs occur at the earlier cognitive stage in language production compared with language comprehension as language production may rely on sensorimotor information from the conceptualization stage (i.e., the first stage of language production) and language comprehension may rely on sensorimotor information after word recognition (i.e., the primary stage of language comprehension).

5.6 Modality Switching Costs in L1

Results of Experiment 1 showed that modality switching costs in L1 were significantly smaller for proficient bilinguals than non-proficient ones. As both proficient and non-proficient bilinguals are native Chinese speakers, their language proficiency of Chinese is supposed to be similar. This difference is most probably caused by the language fluency of L2. It seems that proficient and non-proficient bilinguals process the L1 words in a slightly different way. However, the present study mainly focuses on the difference in switching costs in L2, but not in L1. Therefore, we do not have a prior hypothesis for the difference in modality switching in L1 between proficient and non-proficient bilinguals. As a result, the following discussions are mainly post-hoc.

Before these post-hoc discussions, the term of *language fluency* should be re-defined. Previous studies have found that learning new languages change many types of cognitive abilities (Bialystok, Craik, Klein, & Viswanathan, 2004). In this way, it would be better to treat *language proficient* as a bulk of changes, but not a single factor (Antoniou, 2019). Therefore, the difference on language proficient may include many

DOI: 10.1177/1747021819878089

sub-differences, including different phonetic perception abilities, vocabulary sizes, cognitive control abilities, and so on (Antoniou, Liang, Ettlinger, & Wong, 2015).

Especially, cognitive control gradually changes across the lifespan and could be affected by short-term and long-term experiences, such as learning multiple languages (Craik & Bialystok, 2006; Diamond & Lee, 2011). However, the present discussion on the difference of cognitive control between proficient and non-proficient bilinguals only focuses on the difference that was induced by language fluency, but not other cognitive training or individual differences (Antoniou, 2019).

Results of the present study show that language proficiency may modulate switching costs in L1. This modulation effect may be derived by sub-differences of language proficiency between two groups. One of the most promising candidates is cognitive control. Previous studies have found that learning the second language improved bilinguals' cognitive flexibility, inhibition, and the underlying neural processes (Jiao, Liu, Wang, & Chen, 2019; Sun, Li, Ding, Wang, & Li, 2019; Yang, Ye, Wang, Zhou, & Wu, 2018). Meanwhile, Liu et al. (2019) found that cognitive control is vital for language switching costs. Therefore, proficient bilinguals have better cognitive control ability than non-proficient ones. This ability may be transferred into the modality switching in L1 resulting in smaller modality switching costs.

Unfortunately, the present study did not manipulate cognitive control to test whether the difference in switching costs in L1 between two groups is elicited by cognitive control. Follow-up studies are necessary to test this question.

Some may argue that whether differences in switching costs would appear between two groups even after controlling for individual differences in cognitive abilities.

DOI: 10.1177/1747021819878089

First, we agree that if a cognitive ability assessment is performed, we may find higher cognitive abilities for the proficient group. Second, we are not confident to predict that differences in switching costs between groups would be present even after controlling for individual differences in cognitive abilities. As mentioned above, we only discuss differences in cognitive control that are induced by the language fluency. Following this logic, the prediction of switching costs and cognitive control has two possibilities: (1) if switching costs are mainly elicited by the changes of cognitive control, we predict that differences in switching costs between proficient and non-proficient bilinguals would disappear after controlling for individual differences in cognitive abilities; (2) if switching costs are mainly induced by the changes of other aspects (such as different language processing schemes or competition models between L1 and L2) that are induced by the language fluency, we predict that the differences in switching costs between proficient and non-proficient bilinguals would still appear after controlling for individual differences in cognitive abilities.

As the present study did not focus on the inner mechanisms of switching costs, we did not test whether language proficiency affects cognitive controls (or other aspects) to elicit switching costs. In other words, the present study cannot interpret the role of cognitive control in switching costs, because the present study mainly focuses on the differences of switching costs between proficient and non-proficient bilinguals in L1 and L2. Follow-up studies may further test these questions.

6 Conclusions

The present study adopted a lexical decision (word/pseudo-word judgment) task and a word-naming task to investigate the modality switching costs of proficient and non-

DOI: 10.1177/1747021819878089

proficient Chinese-English bilinguals in the processing of L1 and L2. Results showed that the modality switching costs not only emerged in language comprehension but also in language production in both L1 and L2 processing. These findings suggested that the sensorimotor simulation in language processing occurred spontaneously to some extent. In addition, in L2 processing, the modality switching costs can be found even from non-proficient bilinguals, showing that the sensorimotor simulation in L1 can be transferred to L2 spontaneously. L2 proficiency could modulate the magnitude of the modality switching costs in L2 processing.

7 Acknowledge

This research was funded by the National Social Science Foundation of China (19AYY009).

DOI: 10.1177/1747021819878089

8 References

Ahlberg, D. K., Bischoff, H., Kaup, B., Bryant, D., & Strozyk, J. V. (2017). Grounded cognition: Comparing Language × Space interactions in first language and second language. *Applied Psycholinguistics, 39*, 437-459. doi:10.1017/S014271641700042X

Ansorge, U., Kiefer, M., Khalid, S., Grassl, S., & König, P. (2010). Testing the theory of embodied cognition with subliminal words. *Cognition, 116*, 303-320. doi:10.1016/j.cognition.2010.05.010

Antoniou, M. (2019). The Advantages of Bilingualism Debate. *Annual Review of Linguistics, 5*(1), 395-415. doi:10.1146/annurev-linguistics-011718-011820

Antoniou, M., Liang, E., Ettlinger, M., & Wong, P. C. M. (2015). The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition, 18*(4), 683-695. doi:10.1017/S1366728914000777

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences, 22*, 1177-1187. doi: 10.1017/s0140525x99002149

Barsalou, L. W. (2008a). Cognitive and neural contributions to understanding the conceptual system. *Current Directions in Psychological Science, 17*, 91-95. doi: 10.1111/j.1467-8721.2008.00555.x

Barsalou, L. W. (2008b). Grounded cognition. *Annual Review of Psychology, 59*, 617-645. doi:10.1146/annurev.psych.59.103006.093639

DOI: 10.1177/1747021819878089

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B Biological Sciences, 364*, 1281-1289. doi: 10.1098/rstb.2008.0319

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1 - 48. doi: 10.18637/jss.v067.i01

Baumeister, J. C., Foroni, F., Conrad, M., Rumiati, R. I., & Winkielman, P. (2017). Embodiment and emotional memory in first vs. Second language. *Frontiers in Psychology, 8*. doi:10.3389/fpsyg.2017.00394

Bergen, B., Lau, T.-T. C., Narayan, S., Stojanovic, D., & Wheeler, K. (2010). Body part representations in verbal semantics. *Memory & Cognition, 38*(7), 969-981. doi:10.3758/MC.38.7.969

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, Aging, and Cognitive Control: Evidence from the Simon Task. *Psychology and Aging, 19*(2), 290-303. doi:10.1037/0882-7974.19.2.290

Cai, Z. G., & Connell, L. (2015). Space–time interdependence: Evidence against asymmetric mapping between time and space. *Cognition, 136*, 268-281. doi:10.1016/j.cognition.2014.11.039

Casasanto, D. (2009). Embodiment of abstract concepts: Good and bad in right- and left-handers. *Journal of Experimental Psychology General, 138*, 351-367. doi: 10.1037/a0015854

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*, 3-42. doi: 10.1017/s0142716406060024

Craik, F. I. M., & Bialystok, E. (2006). Cognition through the lifespan: mechanisms of change. *Trends in Cognitive Sciences, 10*(3), 131-138. doi: 10.1016/j.tics.2006.01.007

Dagaev, N. I., & Terushkina, Y. I. (2014). Conceptual knowledge of emotions includes somatosensory component: Evidence from modality-switch cost effect. *Journal of Cognitive Psychology, 26*, 322-332.doi: 10.1080/20445911.2014.892111

Dantzig, S. V., Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2008). Perceptual processing affects conceptual processing. *Cognitive Science: A Multidisciplinary Journal, 32*, 579-590. doi: 10.1080/03640210802035365

De Grauwe, S., Willems, R. M., Rueschemeyer, S.-A., Lemhöfer, K., & Schriefers, H. (2014). Embodied language in first- and second-language speakers: Neural correlates of processing motor verbs. *Neuropsychologia, 56*, 334-349. doi: 10.1016/j.neuropsychologia.2014.02.003

de la Vega, I., de Filippis, M., Lachmair, M., Dudschig, C., & Kaup, B. (2012). Emotional valence and physical space: Limits of interaction. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 375-385. doi:10.1037/a0024979

Declerck, M., & Philipp, A. M. (2015). A sentence to remember: Instructed language switching in sentence production. *Cognition, 137*, 166-173. doi: 10.1016/j.cognition.2015.01.006

Degroot, A. M. B., Dannenburg, L., & Vanhell, J. G. (1994). Forward and backward word translation by bilinguals. *Journal of Memory & Language, 33*, 600-629. doi: 10.1006/jmla.1994.1029

DOI: 10.1177/1747021819878089

Diamond, A., & Lee, K. (2011). Interventions Shown to Aid Executive Function Development in Children 4 to 12 Years Old. *Science, 333*(6045), 959-964. doi:10.1126/science.1204529

Dong, Y., Gui, S., & Macwhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism Language & Cognition, 8*, 221-238. doi: 10.1017/s1366728905002270

Dudschig, C., de la Vega, I., & Kaup, B. (2014). Embodiment and second-language: Automatic activation of motor responses during processing spatially associated L2 words and emotion L2 words in a vertical Stroop paradigm. *Brain & Language, 132*, 14-21. doi: 10.1016/j.bandl.2014.02.002

Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down: object words orient attention to the objects' typical location. *Psychological Science, 19*, 93-97. doi: 10.1111/j.1467-9280.2008.02051.x

Foroni, F. (2015). Do we embody second language? Evidence for ‘partial’ simulation during processing of a second language. *Brain and Cognition*, *99*, 8-16. doi: 10.1016/j.bandc.2015.06.006

Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Fraisse, P. (1967). Latency of different verbal responses to the same stimulus. *Quarterly Journal of Experimental Psychology, 19*, 353-355. doi:10.1080/14640746708400115

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*, 558-565. doi:10.3758/bf03196313

DOI: 10.1177/1747021819878089

Glenberg, A. M., Witt, J. K., & Metcalfe, J. (2013). From the revolution to embodiment: 25 years of cognitive psychology. *Perspectives on Psychological Science, 8*, 573-585. doi:10.1177/1745691613498098

Grauwe, S. D., Willems, R. M., Rueschemeyer, S. A., Lemhöfer, K., & Schriefers, H. (2014). Embodied language in first- and second-language speakers: Neural correlates of processing motor verbs. *Neuropsychologia, 56*, 334-349. doi: 10.1016/j.neuropsychologia.2014.02.003

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*, 301-307. doi: 10.1016/s0896-6273(03)00838-9

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 130-136. doi:10.1037/0096-1523.1.2.130

Jiao, L., Liu, C., Wang, R., & Chen, B. (2019). Working memory demand of a task modulates bilingual advantage in executive functions. *International Journal of Bilingualism*, *23*(1), 102-117. doi:10.1177/1367006917709097

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory & Language, 33*, 149-174. doi: 10.1006/jmla.1994.1008

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The revised hierarchical model: A critical review and assessment. *Bilingualism: Language and Cognition, 13*, 373-381. doi:10.1017/S136672891000009X

Kühne, K., & Gianelli, C. (2019). Is Embodied Cognition Bilingual? Current Evidence and Perspectives of the Embodied Cognition Approach to Bilingual Language Processing. *Frontiers in Psychology*, *10*(108). doi:10.3389/fpsyg.2019.00108

Lachmair, M., Dudschig, C., De, F. M., De, l. V. I., & Kaup, B. (2011). Root versus roof: Automatic activation of location information during word processing. *Psychonomic Bulletin & Review, 18*, 1180-1188. doi: 10.3758/s13423-011-0158-x

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). Multiple perspectives on word production. *Behavioral & Brain Sciences, 22*, 61-69. doi: 10.1017/s0140525x99451775

Li, L., Mo, L., Wang, R., Luo, X., & Chen, Z. (2008). Evidence for long-term cross-language repetition priming in low fluency Chinese–English bilinguals. *Bilingualism: Language and Cognition, 12*, 13-21. doi: 10.1017/s1366728908003453

Liu, C., Jiao, L. U., Wang, Z., Wang, M., Wang, R., & Wu, Y. J. (2019). Symmetries of bilingual language switch costs in conflicting versus non-conflicting contexts. *Bilingualism: Language and Cognition*, *22*(3), 624-636. doi:10.1017/S1366728918000494

Marmolejo-Ramos, F., Cousineau, D., Benites, L., & Maehara, R. (2015). On the efficacy of procedures to normalise Ex-Gaussian distributions. *Frontiers in Psychology, 5*: doi: 10.3389/fpsyg.2014.01548

Marmolejo-Ramos, F., Khatin-Zadeh, O., Yazdani-Fazlabadi, B., Tirado, C., & Sagi, E. (2017). Embodied concept mapping: blending structure-mapping and embodiment theories. *Pragmatics & Cognition*, *24* (2), 164-185.

DOI: 10.1177/1747021819878089

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305-315. doi: 10.1016/j.jml.2017.01.001

Murphy, G. L. (2004). *The big book of concepts*. Cambridge, MA: MIT Press.

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science, 14*, 119-124. doi: 10.1111/1467-9280.t01-1-01429

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2004). Sensorimotor simulations underlie conceptual representations modality-specific effects of prior activation. *Psychonomic Bulletin and Review, 11*, 164-167.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329-347. doi:10.1017/S0140525X12001495

Pylyshyn, Z. W. (1984). *Computation and cognition*. Cambridge, MA: MIT Press.

Shen, M., Xie, J., Liu, W., Lin, W., Chen, Z., Marmolejo-Ramos, F., & Wang, R. (2016). Interplay between the object and its symbol: The size-congruency effect. *Advances in Cognitive Psychology*, *12* (2), 115-129

Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory & Language, 51*, 381-398. doi: 10.1016/j.jml.2004.05.003

Spence, C., Nicholls, M. R., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, *63*(2), 330-336. doi:10.3758/BF03194473

DOI: 10.1177/1747021819878089

Stins, J., Marmolejo-Ramos, F., Wenker, E., Hulzinga, F., & Cañal-Bruland, R. (2017). Words that move us: The effects of sentences on body sway. *Advances in Cognitive Psychology*, *13* (2), 156-165.

Sun, X., Li, L., Ding, G., Wang, R., & Li, P. (2019). Effects of language proficiency on cognitive control: Evidence from resting-state functional connectivity. *Neuropsychologia*, *129*, 263-275. doi:10.1016/j.neuropsychologia.2019.03.020

Thornton, T., Loetscher, T., Yates, M. J., & Nicholls, M. E. (2012). The highs and lows of the interaction between word meaning and space. *Journal of Experimental Psychology: Human Perception & Performance, 39*, 964-973. doi: 10.1037/a0030467

Vallesi, A., Mcintosh, A. R., & Stuss, D. T. (2011). How time modulates spatial responses. *Cortex, 47*, 148-156. doi: 10.1016/j.cortex.2009.09.005

Vermeulen, N., Corneille, O., & Niedenthal P.M. (2008). Sensory load incurs conceptual processing costs. *Cognition, 108*, 287-294 doi:10.1016/j.cognition.2008.09.004

Vermeulen, N., Niedenthal, P. M., & Lumineta, O. (2007). Switching between sensory and affective systems incurs processing costs. *Cognitive Science, 31*, 183-192. doi:10.1080/03640210709336990

Villani, C., Lugli, L., Liuzza, M. T., & Borghi, A. M. (in press). Varieties of abstract concepts and their multiple dimensions. *Language and Cognition*

Vukovic, N., & Shtyrov, Y. (2014). Cortical motor systems are involved in second-language comprehension: Evidence from rapid mu-rhythm desynchronisation. *Neuroimage, 102*, 695-703. doi: 10.1016/j.neuroimage.2014.08.039

DOI: 10.1177/1747021819878089

Wang, R., Fan, X., Liu, C., & Cai, Z. G. (2014). Cognitive control and word recognition speed influence the Stroop effect in bilinguals. *International Journal of Psychology, 51*, 93-101. doi: 10.1002/ijop.12115

Xie, J., Huang, Y., Wang, R., & Liu, W. (2015). Affective valence facilitates spatial detection on vertical axis: Shorter time strengthens effect. *Frontiers in Psychology, 6*. doi:10.3389/fpsyg.2015.00277

Xie, J., Wang, R., & Chang, S. (2014). The mechanism of valence-space metaphors: Erp evidence for affective word processing. *Plos One, 9*. doi:10.1371/journal.pone.0099479

Yang, J., Ye, J., Wang, R., Zhou, K., & Wu, Y. (2018). Bilingual Contexts Modulate the Inhibitory Control Network. *Frontiers in Psychology*, *9*(395). doi:10.3389/fpsyg.2018.00395

Zeelenberg, R., & Pecher, D. (2003). Evidence for long-term cross-language repetition priming in conceptual implicit memory tasks. *Journal of Memory & Language, 49*, 80-94. doi: 10.1016/s0749-596x(03)00020-2

Zwaan, R. A. (1999). Embodied cognition, perceptual symbols, and situation models. *Discourse Processes, 28*, 81-88. doi:10.1080/01638539909545070

DOI: 10.1177/1747021819878089

Figure Captions

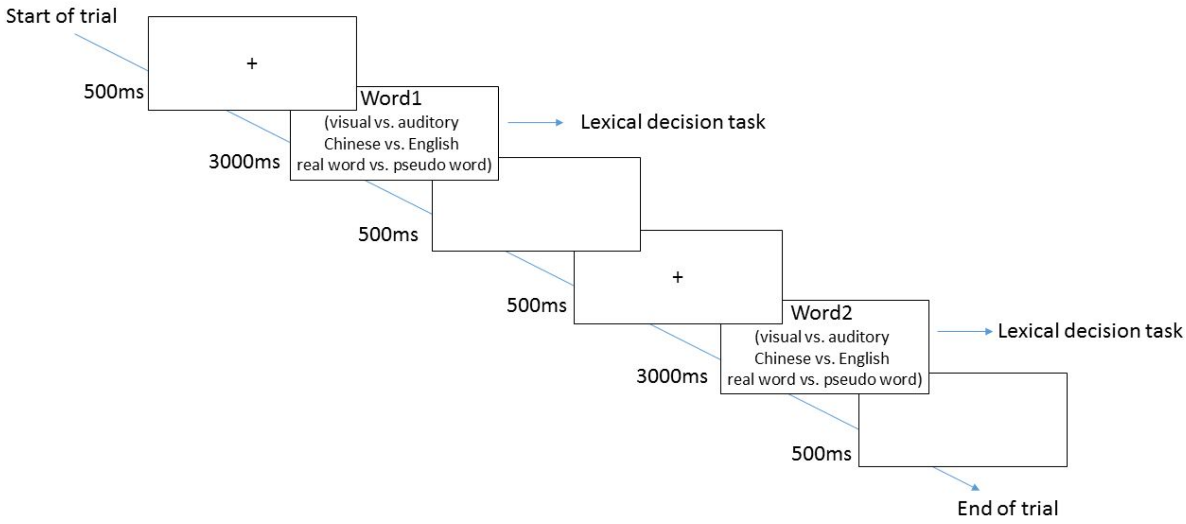
Figure 1. A trial used in Experiment 1.

Figure 2. Violin plots representing the distribution of the RT data in Experiment 1.

Figure 3. Violin plots representing the distribution of the RT data in Experiment 2.

Figure 4. Violin plots representing the distribution of the RT data in Experiment 3.

DOI: 10.1177/1747021819878089



DOI: 10.1177/1747021819878089

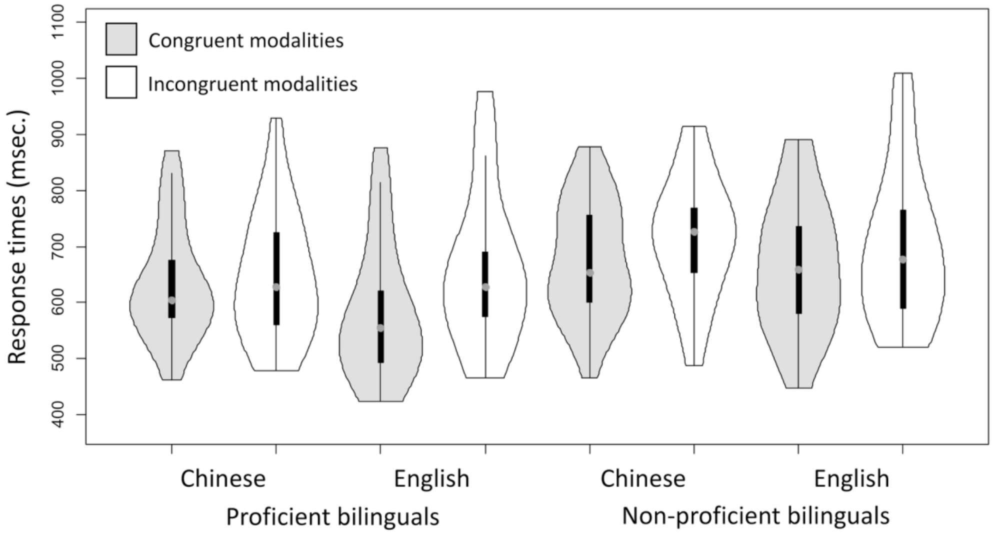


Figure 2. Violin plots representing the distribution of the RT data in Experiment 1.

149x80mm (300 x 300 DPI)

DOI: 10.1177/1747021819878089

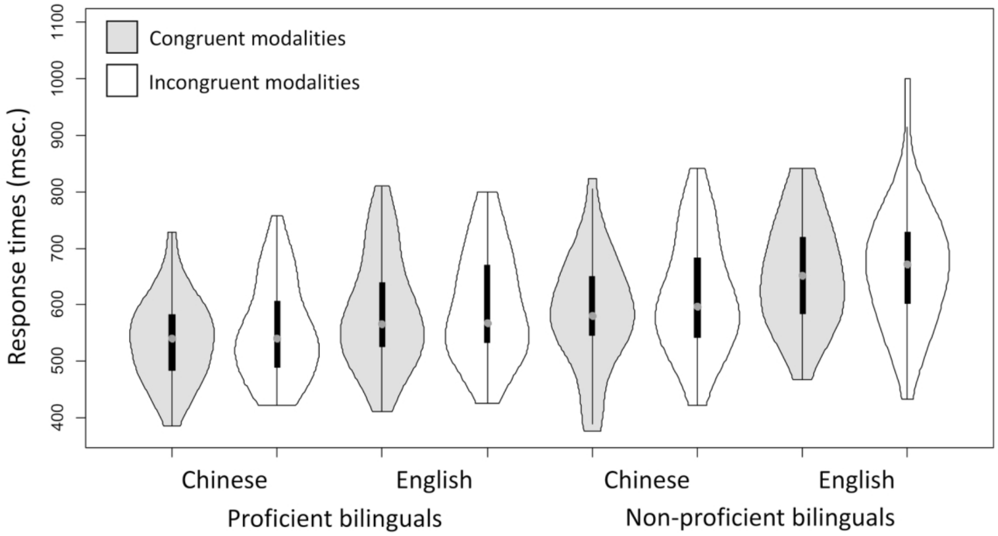


Figure 3. Violin plots representing the distribution of the RT data in Experiment 2.

149x80mm (300 x 300 DPI)

DOI: 10.1177/1747021819878089

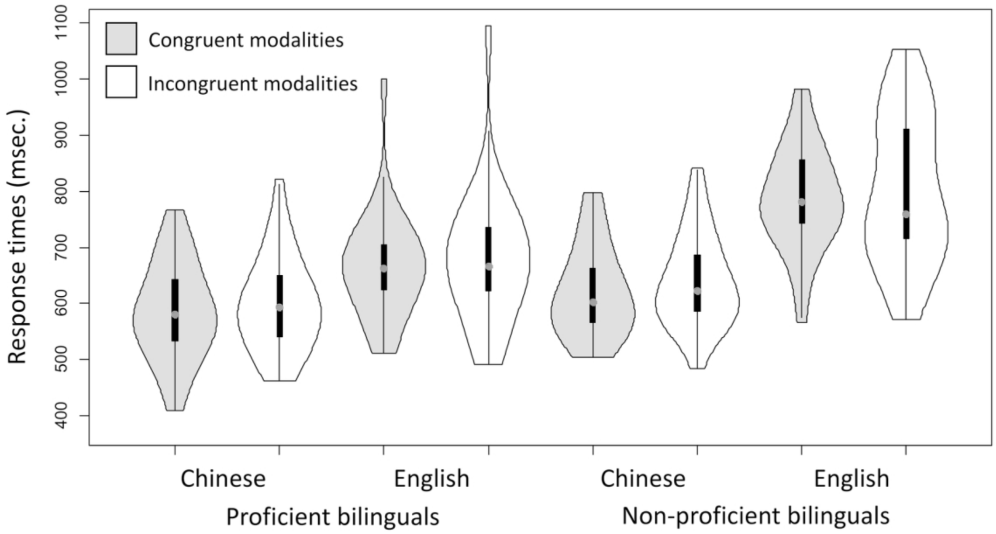


Figure 4. Violin plots representing the distribution of the RT data in Experiment 3.

149x80mm (300 x 300 DPI)

DOI: 10.1177/1747021819878089

Table 1. *Mean and standard deviation of RT (ms) and accuracy (%) in Experiment 1* (*M*±*SD*).

Condition	RT (ms)				Accuracy (%)			
	Chinese		English		Chinese		English	
	Congruent modality	Incongruent modalities	Congruent modality	Incongruent modalities	Congruent modality	Incongruent modalities	Congruent modality	Incongruent modalities
Proficient bilinguals	633.80±100.97	644.35±115.05	582.80±120.11	656.32±134.93	97.14±5.35	97.50±5.18	98.21±3.90	95.00±6.38
Non-proficient bilinguals	674.95±103.14	714.95±107.51	665.13±118.87	699.57±136.12	98.21±3.90	95.71±9.60	93.21±10.20	83.57±10.96

DOI: 10.1177/1747021819878089

Table 2. *Model parameters for the best-fitting model for RT and accuracy in Experiment 1.*

	RT				Accuracy			
Model parameters	Estimate	*SE*	*t*	*p*	Estimate	*SE*	*z*	*p*
Intercept	660.70	15.85	41.70	<.001	3.68	.24	15.25	<.001
L2 proficiency	71.54	26.94	2.66	**.01**	-.75	.29	-2.57	**.01**
Modality congruency	-49.40	18.80	-2.63	**.01**	.54	.41	1.32	.19
Word language	-14.57	21.11	-.69	.49	-.72	.40	-1.79	.07
L2 proficiency × modality congruency	-4.82	17.30	-.28	.78	.60	.49	1.22	.22
L2 proficiency × word language	9.86	25.84	.38	.70	-1.46	.49	-2.96	**<.01**
Modality congruency × word language	-31.90	37.24	-.86	.40	.36	.81	.45	.66
L2 proficiency × modality congruency × word language	74.93	33.02	2.27	**.03**	-.94	.98	-.96	.34

DOI: 10.1177/1747021819878089

Table 3. *Mean and standard deviation of RT (ms) and accuracy (%) in Experiment 2* (*M*±*SD*).

Condition	RT (ms)				Accuracy (%)			
	Chinese		English		Chinese		English	
	Congruent modality	Incongruent modalities	Congruent modality	Incongruent modalities	Congruent modality	Incongruent modalities	Congruent modality	Incongruent modalities
Proficient bilinguals	537.37± 74.16	551.82± 85.46	586.59± 97.17	596.90± 96.02	97.50± 5.08	98.44± 3.69	97.50± 5.08	96.88± 6.44
Non-proficient bilinguals	586.04± 98.67	615.06± 101.58	654.80± 94.90	666.97± 109.10	97.50± 5.68	99.37± 2.46	93.75± 6.60	93.13± 7.80

DOI: 10.1177/1747021819878089

Table 4. *Model parameters for the best-fitting model for RT and accuracy in Experiment 2.*

	RT				Accuracy			
Model parameters	Estimate	*SE*	*t*	*p*	Estimate	*SE*	*z*	*p*
Intercept	604.16	15.52	38.93	<.001	4.58	.37	12.29	<.001
L2 proficiency	72.44	20.43	3.55	**<.001**	-.61	.49	-1.24	.22
Modality congruency	-18.08	5.52	-3.28	**.001**	-.27	.28	-.98	.33
Word language	57.90	25.02	2.31	**.03**	-1.46	.59	-2.48	**.01**
L2 proficiency × modality congruency	-8.33	11.03	-.76	.45	-.42	.55	-.77	.44
L2 proficiency × word language	16.19	17.87	.91	.37	-1.14	.65	-1.75	.08
Modality congruency × word language	14.03	11.04	1.27	.20	.88	.55	1.61	.11
L2 proficiency × modality congruency × word language	21.92	22.06	.99	.32	.31	1.09	.29	.77

DOI: 10.1177/1747021819878089

Table 5. *Mean and standard deviation of RTs (ms) and accuracy (%) in Experiment 3* (*M*±SD).

Condition	RT (ms)				Accuracy (%)			
	Chinese		English		Chinese		English	
	Congruent modality	Incongruent modalities	Congruent modality	Incongruent modalities	Congruent modality	Incongruent modalities	Congruent modality	Incongruent modalities
Proficient bilinguals	585.26± 86.31	599.06± 84.82	666.21± 94.81	677.80± 119.61	98.57± 3.50	98.93± 3.15	98.57± 4.48	99.29± 2.62
Non-proficient bilinguals	621.31± 81.46	638.91± 82.61	787.40± 92.83	802.08± 131.19	97.14± 5.34	96.79± 5.48	97.14± 4.60	95.71± 7.90

DOI: 10.1177/1747021819878089

Table 6. *Model parameters for the best-fitting model for RTs in Experiment 3.*

Model parameters	Estimate	*SE*	*t*	*p*
Intercept	679.97	18.49	36.78	<.001
L2 proficiency	93.97	23.51	4.00	**<.001**
Modality congruency	-19.84	5.36	-3.70	**<.001**
Word language	129.40	31.48	4.11	**<.001**
L2 proficiency × modality congruency	-23.33	10.72	-2. 18	**.03**
L2 proficiency × word language	104.53	26.57	3.94	**<.001**
Modality congruency × word language	-14.83	10.66	-1.39	.17
L2 proficiency × modality congruency × word language	-30.20	21.31	-1.42	.16

DOI: 10.1177/1747021819878089