# Limitations of translation activation in masked priming: Behavioural evidence from Chinese-English bilinguals and computational modelling

**Yun Wen[1, 2]**

**Walter J. B. van Heuven[1]**

**[1] School of Psychology, University of Nottingham, Nottingham, UK**

**[2] Laboratoire de Psychologie Cognitive, Aix-Marseille Université and Centre National de la Recherche Scientifique, Marseille, France**

Running head: Limitations of Translation Activation in Masked Priming

**Corresponding author:**

Yun Wen

Laboratoire de Psychologie Cognitive

CNRS AMU LPC UMR 7290, Campus St Charles

3, Place Victor Hugo

13331 Marseille, France

E-mail : yun.wen@univ-amu.fr

**Abstract**

Electrophysiological and behavioural evidence suggests that Chinese translations of English words are automatically activated when Chinese-English bilinguals read English words (e.g., Thierry & Wu, 2007; Wu & Thierry, 2010; Zhang et al., 2011). The present study investigated the impact of translation activation in three behavioural experiments with in total 118 Chinese-English bilinguals. First, we investigated whether Chinese phonology was the source of the effects of Chinese character repetition in the Chinese translations of English masked primes and targets (hidden repetition priming) observed in Zhang et al's (2011), and whether these hidden repetition priming effects were affected by Chinese morpheme complexity and prime duration. However, we failed to find any evidence of hidden repetition priming. An exact replication of Zhang et al. (2011) was conducted next, which again provided no evidence for hidden repetition priming. However, cross-language priming data collected with the same group of participants did reveal masked translation priming and crucially Chinese character repetition priming with masked Chinese primes and English targets (partially hidden repetition priming), indicating that the activation of Chinese translations in the masked priming paradigm is limited to English target words. Computational modeling work provided further support that translation form activation is limited to target words in the masked priming paradigm.

**Keywords**

**INTRODUCTION**

Bilinguals have the unique ability to translate words between their languages. Although translation seems a deliberate and conscious process, recent research has shown that first language (L1) translation equivalents are automatically activated during second language (L2) word reading (e.g., Meade, Midgley, Sevcikova Sehyr, Holcomb, & Emmorey, 2017; Morford, Wilkinson, Villwock, Piñar, & Kroll, 2011; Thierry & Wu, 2007; Wu, Cristino, Leek, & Thierry, 2013; Wu & Thierry, 2010, 2012a, 2012b; Zhang, van Heuven, & Conklin, 2011). In an event-related potentials (ERP) study, Thierry and Wu (2007) presented proficient Chinese-English bilinguals with pairs of English words and asked them to judge whether these word pairs were related in meaning or not. Unknown to the participants, the Chinese translation of the critical English word pairs had a repeated character, e.g., train [火车] - ham [火腿]. The logic of this elegant design is that, when bilinguals performed differently to the English word pairs with a hidden repeated character compared to control pairs without hidden repeated characters, the Chinese translations of the English words must have been activated. A reduction of the N400 amplitude was found for English word pairs with repeated characters in their Chinese translations. Therefore, the ERP data provided evidence for a hidden priming repetition effect (hereafter the term hidden repetition is used to refer to the repetition in the Chinese translations of the English word pairs).

In their follow-up study, Wu and Thierry (2010) further investigated whether Chinese phonology and/or orthography was activated. Using the same paradigm, Chinese translations of critical English pairs either shared a homophone or a homograph. For example, *experience* [经验, /**Jing1**Yan4/] - *surprise* [惊讶, /**Jing1**Ya4/] created a pure hidden phonological repetition, whereas the hidden repetition of *accountant* [会计, /Kuai4Ji4/] - *conference* [会议, /Hui4Yi4/] was only in the orthographic forms. The ERP data showed a reduced N400 for the

hidden phonological repetition but not for the hidden orthographic repetition, suggesting that Chinese-English bilinguals activate the phonology rather than the orthography of the Chinese translations. This Chinese phonological activation during English word reading was successfully replicated in another electrophysiological study by Wu and Thierry (2012a) which manipulated the emotional valence of prime words. Furthermore, additional evidence for Chinese phonological activation during English word reading came from Wu and Thierry's (2012b) study that involved a non-linguistic task. In this study, Chinese-English bilinguals were presented with shapes (e.g., square or circle) mixed with English words. The participants' task was to differentiate circles and squares by pressing buttons and to ignore English words. Importantly, critical English words had phonological overlap with the words *circle* or *square* when translated into Chinese (e.g., the first character of the Chinese translation of *reason* is a homophone of the Chinese translation of *circle*). When comparing these critical words with hidden phonological overlap to the control words, an increased N200 amplitude was observed. Because a more negative-going N200 was assumed to index inhibition, these findings revealed that the activation of Chinese phonology of translations from critical English words caused interference in bilinguals. Adapting this novel paradigm in an eye-tracking study, Wu et al. (2013) also found the evidence of phonological activation of Chinese translations during English words reading in Chinese-English bilinguals. Taken together, these studies provided a strong case for the co-activation of translation equivalents.

It is important to point out that the electrophysiological and eye-tracking studies by Wu and Thierry always involved processing of *visible* English words. Furthermore, the evidence for the activation of translations was found in ERP and eye-tracking data but not always in behavioural responses. To investigate whether Chinese translations are activated when English words are *subliminally* presented, Zhang et al. (2011) combined the masked priming technique with the hidden repetition paradigm developed by Thierry and Wu (Thierry & Wu,

2007; Wu & Thierry, 2010). In Zhang's et al. (2011) behavioural masked priming study, Chinese-English bilinguals were asked to perform a lexical decision task with English target words, which were preceded by English primes presented for 59 ms. Faster responses were found to targets when the Chinese translation of the prime was the first character in the Chinese translation of the target (e.g., *east* [东] - *thing* [东西]), but not when the Chinese translation of the prime was the second character in the Chinese translation of the target (e.g., *west* [西] - *thing* [东西]). These results provided behavioural evidence for the activation of Chinese translations of the English prime and target. Although the findings of Zhang et al. (2011) support the idea of fast and automatic translation of primes and targets, it is unclear whether this behavioural masked hidden repetition priming is also driven by Chinese phonology.

The hidden repetition priming observed in English originates from the activation of L1 (Chinese) and can therefore only occur when both the L2 (English) prime and L2 target are translated into L1 (Zhang et al., 2011). The activation of translations is consistent with the assumption of non-selective lexical access (e.g., Dijkstra, 2005; Dijkstra & van Heuven, 2002; van Heuven, Dijkstra, & Grainger, 1998). However, the L2 primes in Zhang et al., (2011) were presented very briefly (59 ms). Therefore, it is remarkable that the L1 translations of the L2 primes were activated because masked priming studies using non-cognate translation equivalents with similar prime durations have found strong L1-L2 (L1 primes and L2 targets) translation priming effects, whereas L2-L1 (L2 primes and L1 targets) translation priming effects have been found to be much smaller (but still significant, for a recent meta-analysis see Wen & van Heuven, 2017a). This translation asymmetry has been explained in terms of different strengths between lexical and semantic links for L1 and L2 (Kroll & Tokowicz, 2001) or slower activation of L2 representations (Allen, Conklin, & van Heuven, 2015), but it can also be attributed to a representational asymmetry between L1 and

L2 words, with L1 words connected to more semantic features than L2 words (Finkbeiner, Forster, Nicol, & Nakamura, 2004; Schoonbaert, Duyck, Brysbaert, & Hartsuiker, 2009). In all cases, it seems unlikely that a translation asymmetry would occur if L2 primes would quickly activate the L1 translations as in Zhang et al. (2011). The activation of L1 translations by briefly presented L2 primes has not been observed in masked translation priming studies using ERP, because N250 effects (associated with translation form activation) were found for L1-L2 priming, but not for L2-L1 priming (e.g., Hoshino, Midgley, Holcomb, & Grainger, 2010; Midgley, Holcomb, & Grainger, 2009). Thus, the findings in the literature seem to indicate that masked L2 primes do not activate their L1 translation forms. It is also remarkable that the 75 ms hidden repetition priming effect observed in Experiment 1 of Zhang's et al. (2011), which required the translation of L2 primes into L1, is similar in size to the L1-L2 translation priming effect reported in other masked priming studies with Chinese-English bilinguals (i.e., a mean priming effect of 69 ms from 4 experiments: Chen, Zhou, Gao, & Dunlap, 2014, Experiment 2; Jiang, 1999, Experiment 1; Xia & Andrews, 2015, Experiment 1B and 2B). Thus, it is crucial to further investigate the hidden repetition priming in the masked priming paradigm.

The present study aimed to further investigate hidden repetition priming in the masked priming lexical decision task and to determine the source of the hidden repetition priming effects observed by Zhang et al. (2011). To investigate whether Chinese phonology or orthography drives hidden repetition priming, Experiment 1 compared English word pairs with repeated characters (repeated orthography and phonology) in their Chinese translations and English word pairs with only repeated phonology in their Chinese translations. If hidden repetition priming is driven only by Chinese phonology, a priming effect would be expected for English word pairs with repeated phonology when translated into Chinese. This priming effect would be comparable to the priming effect for English word pairs with repeated

characters in their Chinese translations. Alternatively, if hidden repetition priming is not driven by Chinese phonology, a priming effect would *only* be expected for English word pairs that have translations with repeated Chinese characters. To test whether the results favoured the null hypothesis (no priming effect) or the alternative hypothesis (a priming effect), the Bayes factor (BF) was also calculated by the BayesFactor package in R (Morey & Rouder, 2016). A BF smaller than 0.3 provides support for the null hypothesis and a BF larger than 3 provides support for the alternative hypothesis, whereas a BF between 0.3 and 3 supports neither hypothesis, which would indicate insensitive data (Dienes, 2014).

**EXPERIMENT 1**

**Methods**

*Participants*

33 Chinese-English bilinguals participated in Experiment 1. One participant was excluded because of high error rates in the experimental trials (> 25%). In this experiment as well as in the following experiments, the bilingual participants were all native Mandarin Chinese speakers who learnt Chinese from birth. Participants were undergraduate or postgraduate university students studying in Nottingham. They met the minimum English language entry requirements to study at the University of Nottingham (IELTS 6.0 for undergraduates and 6.5 for postgraduates). The details of participants' language background are summarized in Table 1

Table 1 Summary of participants' language background data from 3 experiments

|  | Mean (SD) | | |
|  | Experiment 1 | Experiment 2 | Experiment 3 |
|---|---|---|---|
| Age (years) | 22.5 (2.7) | 21.2 (1.3) | 22.8 (2.2) |
| Age exposed to formal English education | 9.3 (2.2) | 8.2 (1.8) | 8.4 (2.8) |
| Time studies English (years) | 13.2 (2.4) | 13.0 (2.1) | 14.1 (3.1) |
| English immersion experience (months) | 6.2 (10.9) | 4.0 (3.7) | 9.0 (12.3) |
| Self-rated English Reading ability | 4.9 (1.0) | 5.3 (0.8) | 5.1 (0.8) |
| LexTALE test score | 62.7 (8.0) | 62.3 (9.4) | 63.5 (8.3) |

*Note*. LexTALE (Lemhöfer & Broersma, 2012) is a quick and valid English vocabulary test; subjective reading ability were rated on a 7-point scale (1 = very poor, 7 = native-like). Results of Mann-Whitney Test showed that there was no significant difference in the self-ratings between participants from across 3 experiments ($p > 0.25$) and independent t-tests on their LexTALE scores also revealed no significant difference between the participant groups ($p > 0.25$).

*Materials and design*

Stimuli were selected from an English-Chinese translation norming study (Wen & van Heuven, 2017b). The stimulus material consisted of two sets of 40 English target words. In the first set, each of the 40 target words was preceded by either an English word whose two-character Chinese translation shared the first character with the Chinese translation of the target word (hidden phonological and orthographic repetition prime), or by an English control word matched in length and word frequency. In the second set, each of the 40 target words was preceded by either an English word in which the first character in the Chinese translation

was the homophone of the first character in the Chinese translations of the target word (hidden phonological repetition prime), or by an English control word matched in length and word frequency. Importantly, all English prime - target word pairs were orthographically, phonologically, and semantically unrelated in English. Examples of the stimuli are presented in Table 2. For the English lexical decision task, eighty English pseudowords were selected from the British Lexicon Project (Keuleers, Lacey, Rastle, & Brysbaert, 2012). These pseudowords were matched with the target words in length and the first letter. Pseudowords were preceded by 80 unrelated English words. Two counterbalanced lists were created and participants were randomly assigned to one of the lists.

Table 2 Experimental design, mean lexical characteristics of prime and target words, and stimulus examples of Experiment 1

| | *Phonological & Orthographic Repetition* | | | *Phonological Repetition* | | |
|---|---|---|---|---|---|---|
| | *PO repetition* | *PO control* | *Target* | *P repetition* | *P control* | *Target* |
| English | career | cancer | fact | winter | wealth | thing |
| Chinese | 事业 | 癌症 | 事实 | 冬天 | 财富 | 东西 |
| Pinyin | *Shi4*Ye4 | Ai2Zheng4 | *Shi4*Shi2 | *Dong1*Tian1 | Cai2Fu4 | *Dong1*Xi1 |
| Word Length | 6.23 | 6.18 | 5.78 | 6.13 | 6.10 | 5.35 |
| English Frequency | 4.44 | 4.55 | 4.65 | 4.52 | 4.52 | 4.76 |
| Concreteness | 3.99 | 3.51 | 3.79 | 3.48 | 3.98 | 3.97 |

*Note*. English frequencies are Zipf values from SUBTLEX-UK (van Heuven, Mandera, Keuleers, & Brysbaert, 2014). Concreteness scores are based on Brysbaert, Warriner, and Kuperman (2014). Pinyin is a Romanisation of the Chinese pronunciation.

*Procedure*

Participants were tested individually in a sound-attenuated experimental room. The stimuli were presented on a monitor using DMDX Display Software (Forster & Forster, 2003). Each participant received a unique random presentation order. The procedure was identical to Zhang et al. (2011). Each trial began with a forward mask ($@#$£@£%) presented for 500 ms. Following this mask was a lowercase English word presented for 59 ms, followed by a backward mask (%$%£@£$#) presented for 24 ms. Immediately following the backward mask the target, an uppercase letter string, was presented. Participants were not informed of the presence of the primes and were asked to decide whether the uppercase letter string was an English word or not by pressing buttons on a Cedrus response box attached to the computer.[1] They were instructed to perform the task as quickly and as accurately as possible. The uppercase string remained on the screen up to a maximum of 3000 ms or until the participants responded. The inter-trial interval was set at 1000 ms. The experiment was approved by the Ethics Committee of the School of Psychology at the University of Nottingham. All participants signed a consent form before the experiment started.

**Results and discussion**

The mean reaction times and error rates of Experiment 1 are presented in Table 3. Incorrect responses for the critical stimuli were discarded (7.7%, error rates for the pseudowords: 12.8%). Reaction times less than 300 ms or more than 2,500 ms were discarded as outliers (0.3%). Because the two critical priming conditions (i.e., hidden phonological and orthographical repetition prime and hidden phonological repetition prime) used different target words that were not matched on lexical properties, and each of the prime conditions was matched with different controls, priming effects were analysed separately (see Guo, Misra, Tam, & Kroll, 2012 for a similar approach). Thus, paired by-subject ($t_1$) and by-item ($t_2$) t-tests were conducted separately for both priming conditions and their controls using

reaction times and error rates. The results revealed no significant differences between the hidden orthographic and phonological repetition condition and the control condition in reaction times, all $t$s < 1, $p$s > .25, all BFs < 0.3, and error rates, $t_1$ (31) = 1.735, $p$ = .093, BF = 0.71 and $t_2$ (39) = 1.397, $p$ = .170, BF = 0.42. For the hidden phonological repetition prime conditions, the analyses of error rates and reaction times revealed no significant differences, all $t$s < 1, $p$s > .25, all BFs < 0.3.

Table 3 Mean reaction times and error rates (with SD in brackets) of Experiment 1

| | *Phonological & Orthographic Repetition* | | *Phonological Repetition* | |
| --- | --- | --- | --- | --- |
| | *PO repetition* | *PO control* | *P repetition* | *P control* |
| RT (ms) | 946 | 963 | 913 | 914 |
| | (211) | (231) | (195) | (204) |
| Error Rates (%) | 10.87 | 8.63 | 5.47 | 6.10 |
| | (9.41) | (8.54) | (5.73) | (6.31) |

In contrast to Zhang et al. (2011), no evidence for hidden repetition priming was found. Interestingly, ERP studies using the hidden repetition paradigm also reported no significant hidden repetition priming effects in their behavioural data (Thierry & Wu, 2007; Wu & Thierry, 2010, 2012a; cf. Thierry & Wu, 2004), although primes and targets were both clearly visible in these studies (primes presented for 500 ms in a semantic relatedness judgment task). The difference between the findings of Zhang et al. (2011) with a masked priming paradigm and the ERP studies using a semantic relatedness judgment task (unmasked) paradigm is puzzling, because the priming effects are generally stronger in an unmasked paradigm than in a masked paradigm (e.g., Brown & Hagoort, 1993; de Groot & Nas, 1991). However, a possible explanation for the behavioural effects observed in Zhang et al. (2011) is the morphological characteristics of the Chinese translations. Zhang et al. (2011) used primes

which can be translated into one-character Chinese words (e.g., east - 东), whereas the Chinese translations of the English pairs are two-character words in the current experiment and in previous ERP studies (e.g., train - 火车). To investigate this, Experiment 2 examined the impact of morphological complexity on hidden repetition priming (i.e., either the Chinese translation of the prime is a two-character or a one-character word). If the morphological complexity of the Chinese translations does impact hidden repetition priming, effects would only be observed when English primes are translated into one-character Chinese words. Furthermore, it is important to examine whether hidden repetition priming would get stronger with a longer prime duration because priming effects are often more salient when primes are presented longer (Chen et al., 2014; Grainger & Frenck-Mestre, 1998). Therefore, the prime duration in the next experiment was also manipulated (59 vs. 118 ms).

**EXPERIMENT 2**

**Methods**

*Participants*

34 Chinese-English bilinguals participated in Experiment 2. Two participants were excluded, one because of high error rates in the experimental trials (>25%) and the other because of a failure to follow the task instructions. The details of participants' language background data are summarized in Table 1.

*Materials and design*

The stimulus material consisted of two sets of forty English words used as targets. The first set of 40 words were taken from Experiment 1 of Zhang et al. (2011). These words were preceded by either English words whose one-character Chinese translations were the first characters of the Chinese translations of the target words, or by English control words

matched in length and word frequency. Using the data from an English-Chinese translation norming study (Wen & van Heuven, 2017b), we selected a second set of 40 English target words. These words were preceded by either an unrelated English word whose two-character Chinese translations shared the first character with the Chinese translation of the target word, or by an unrelated English control word matched in length and word frequency. Examples of stimuli are presented in Table 4. The duration of the primes (59 ms vs. 118 ms) was further manipulated in two counterbalanced blocks. Eighty English pseudowords were selected from the British Lexicon Project (Keuleers et al., 2012). These pseudowords were matched with the target words in length and the first letter. Similar to Experiment 1, pseudowords were preceded by 80 English words. Four counterbalanced lists were created for both sets of target words to ensure that in each list each target was viewed once and primed with all 4 experimental conditions (2 prime types × 2 prime durations) across lists. Participants were randomly assigned to one of the lists.

Table 4 Experimental design, mean lexical characteristics of prime and target words, and stimulus examples of Experiment 2

|  | *Hidden One-Character Prime* | | | *Hidden Two-Character Prime* | | |
|---|---|---|---|---|---|---|
|  | *Repetition* | *Control* | *Target* | *Repetition* | *Control* | *Target* |
| Word | water | night | level | career | cancer | fact |
| Chinese | 水 | 夜 | 水平 | 事业 | 癌症 | 事实 |
| Word Length | 5 | 5 | 6.3 | 6.23 | 6.18 | 5.78 |
| English Frequency | 5.16 | 5.12 | 4.72 | 4.44 | 4.55 | 4.65 |
| Concreteness | 3.65 | 3.71 | 3.28 | 3.60 | 4.06 | 3.85 |

*Note*. English frequencies are Zipf values from SUBTLEX-UK (van Heuven et al., 2014).

Concreteness scores are based on Brysbaert et al. (2014).

*Procedure*

The procedure was similar to Experiment 1 except for two aspects. First, the prime duration was either 59 ms or 118 ms, this was manipulated between blocks and the order of the blocks was counterbalanced across participants. Second, the uppercase string remained on the screen up to a maximum of 4000 ms instead of 3000 ms.

**Results and discussion**

The mean error rates and reaction times for Experiment 2 are presented in Table 5. Incorrect responses for the critical stimuli were discarded (9.5%, error rates for the pseudowords: 12.2%). Reaction times less than 300 ms or more than 3,000 ms were discarded as outliers (0.7%).[2] Because the two prime conditions (i.e., Hidden One-character Prime and Hidden Two-character Prime) had different target words that were not matched on lexical properties, and each of the hidden repetition prime was matched with different controls, priming effects were as in Experiment 2 analysed separately. For both Hidden One-character Prime and Hidden Two-character Prime condition, a 2 (control vs. hidden repetition) × 2 (59 ms vs. 118 ms) repeated measures ANOVA was separately run using reaction times and error rates across subjects ($F_1$) and across items ($F_2$). For Hidden One-character Prime, reaction times revealed no significant effects of prime duration, hidden repetition priming, and no significant interaction, all $Fs < 1$, $ps > .25$, BFs $< 0.3$ expect for prime duration in the by-subject analysis, BF $= 0.31$. The error analysis revealed also no significant effects of prime duration, $F_1 (1, 31) = 1.4$, $p = .246$, BF $= 0.47$, $F_2 (1, 39) = 1.414$, $p = .242$, BF $= 0.4$, and hidden repetition priming, and no significant interaction, all $Fs < 1$, $ps > .25$, BFs $< 0.3$. For the Hidden Two-character condition, reaction times also revealed no significant effects of prime duration, $Fs < 1$, $ps > .25$, BFs $< 0.3$, hidden repetition priming, $F_1 (1, 31) < 1$, $p > .25$, BF $<0.3$, $F_2 (1, 39) = 1.05$, $p >.25$, BF $< 0.3$, and no interaction, $Fs < 1$, $ps > .25$, BFs $< 0.3$. Similarly, the error analysis revealed no significant effects of prime

duration, hidden repetition priming, $F$s < 1, $p$s > .25, BFs < 0.3, and no significant interaction, $F_1$ (1, 31) < 1, $p$ > .25, BF < 0.3, $F_2$ (1, 39) = 1.372, $p$ =.249, BF < 0.3.

Table 5 Mean reaction times and error rates (with SD in brackets) in Experiment 2

| | *Hidden One-Character Prime* | | | | *Hidden Two-Character Prime* | | | |
|---|---|---|---|---|---|---|---|---|
| Prime Duration | *59 ms* | | *118 ms* | | *59 ms* | | *118 ms* | |
| Priming Condition | *Repetition* | *Control* | *Repetition* | *Control* | *Repetition* | *Control* | *Repetition* | *Control* |
| RT (ms) | 1093 (311) | 1068 (281) | 1047 (270) | 1055 (228) | 1099 (249) | 1096 (251) | 1084 (285) | 1091 (275) |
| Error Rate (%) | 6.0 (6.2) | 6.6 (7.5) | 8.2 (6.9) | 8.4 (11.1) | 12.0 (11.5) | 12.7 (10.9) | 12.3 (11.6) | 10.0 (8.4) |

To summarize, no evidence was found for hidden repetition priming. Thus, we failed to replicate the findings of Zhang et al. (2011) with one-character primes and 59 ms prime duration. To understand these unexpected findings, it is crucial to investigate whether the hidden repetition priming findings can be replicated at all. Thus, an exact replication of Zhang et al. (2011) is needed because the design of the current experiment was not identical to the original study. To ensure that this replication has enough statistical power to detect an effect, the number of participants needed for this replication experiment was determined by a power analysis using the effect size of Zhang et al. (2011) (d = 0.349, power = 0.8, one-tailed paired t-test). The results of this power analysis indicated that 52 participants were needed.

To investigate whether a possible null effect is not due to technical or procedural issues and to confirm that participants are processing the primes, we included two additional tasks in this experiment. Thus, the experiment consisted of three tasks, which were conducted in a fixed order so that the first task was an exact replication of Zhang et al. (2011). The second

task was an English lexical decision task with Chinese primes and English targets (cross-language masked priming) using the same prime duration as in the first task (59 ms). Because robust L1-L2 translation priming effects have been reported in numerous studies (e.g., Duñabeitia, Dimitropoulou, Uribe-Etxebarria, Laka, & Carreiras, 2010; Gollan, Forster, & Frost, 1997; Hoshino et al., 2010; Wang & Forster, 2010), we expected significant masked translation priming effects. These would provide evidence that participants in this study processed the primes, ruling out the possibility that the absence of hidden repetition priming effect in the first task is due to the participants' failure to process the primes or to any experimental issue in terms of the hardware and software used. Furthermore, in the second task we investigated whether Chinese translations of English target words are activated by using Chinese prime words that were unrelated in meaning with the English target words but that shared the first character of the Chinese translation of the English target words (e.g., prime 事业, meaning "career", and target FACT, which Chinese translation is 事实, repeated character 事). From this point onwards we will refer to this type of repetition as *partially* hidden repetition priming. Only when the Chinese translations of English target words are activated, the Chinese character repetition is able to impact participants' response times. Thus, a hidden repetition priming effect in task 1 (English primes and targets) can occur only when English primes and targets are both translated into Chinese, whereas a partially hidden repetition priming effect in task 2 (Chinese primes and English targets) can only occur when English target words are translated into Chinese. Therefore, if we find a partially hidden repetition priming effect in the second task but not a hidden repetition priming effect in the first task, it would indicate that the activation of Chinese translations is restricted to English target words in the masked priming lexical decision task. Furthermore, a partially hidden repetition priming effect, if observed, should be smaller than the L1-L2 translation priming effect because the partially hidden repetition priming effect originates from a form overlap,

whereas the L1-L2 translation priming effect is driven by the semantic overlap (Hoshino et al., 2010; Midgley et al., 2009; Schoonbaert, Holcomb, Grainger, & Hartsuiker, 2011).

Potentially, we might not observe any hidden or partially hidden repetition priming effects at all in the first two tasks. To investigate whether this would be due to the short (masked) presentation of the prime (59 ms), a translation recognition task was included as the final (third) task of this experiment. A translation recognition task requires explicit processing of the prime (Chinese words) and target (English words) in order to decide whether the prime and target are translation equivalents. Therefore, the prime is presented for 500 ms in order for participants to consciously translate the prime. The design of this experiment and the stimuli included were the same as for the second task. Previous studies using translation recognition tasks have reported significant through-translation interference effects when primes and targets were not translation equivalents but contained form similarities (e.g., Ferré, Sánchez-Casas, & Guasch, 2006; Guasch, Sánchez-Casas, Ferré, & García-Albea, 2008; Guo et al., 2012; Ma, Chen, Guo, & Kroll, 2017; Moldovan, Demestre, Ferré, & Sánchez-Casas, 2016; Moldovan, Sánchez-Casas, Demestre, & Ferré, 2012), which indicates that translation equivalents were activated. For example, in Guo et al. (2012), Chinese-English bilinguals conducted an English-Chinese (L2-L1) translation recognition task. Participants were less accurate and slower in responding when the Chinese targets (e.g., 塘) had a similar pronunciation and form with Chinese translations of the English primes (e.g., *sugar* whose Chinese transition is 糖), suggesting that L1 translations of L2 words were activated. Therefore, we expected that partially hidden repetition would lead to an interference effect in the translation recognition task.

**EXPERIMENT 3**

**Methods**

*Participants*

57 Chinese-English bilinguals participated in Experiment 3. Five participants were excluded from the analysis due to high error rates in the experimental trials (> 25%). Details of the participants' language background data in this experiment are summarized in Table 1.

*Materials and design*

The experiment consisted of three tasks that were presented in a fixed order so that task 1, the critical replication of the first experiment of Zhang et al. (2011), was conducted first. The first two tasks were masked priming lexical decision tasks, whereas the final task was a translation recognition task. The stimuli of the first task were identical to Experiment 1 of Zhang et al. (2011). Thus, 40 English target words were preceded by two critical primes and two matched control primes. The two critical prime conditions consisted of English words whose Chinese translations were identical to the first character of the Chinese translation of the English target word, and an English words whose Chinese translations were identical to the second character of the Chinese translations of the English target words. These two prime conditions were matched with two control conditions that consisted of English primes matched in length and English frequency with the coresponding critical primes. Examples of stimuli are presented in Table 6. Four counterbalanced lists were created and participants were randomly assigned to one of the lists.

Table 6 Experimental design, mean lexical characteristics of prime and target words, and stimulus examples of Experiment 3 task 1

| | First Character | | | Second Character | | |
|---|---|---|---|---|---|---|
| | *Repetition* | *Control* | *Target* | *Repetition* | *Control* | *Target* |
| Word | east | wall | thing | west | town | thing |
| Chinese Translation | 东 | 墙 | 东西 | 西 | 镇 | 东西 |
| Word Length | 5.0 | 5.0 | 6.3 | 4.9 | 4.9 | 6.3 |
| English Frequency | 5.16 | 5.12 | 4.72 | 5.13 | 5.10 | 4.72 |
| Concreteness | 3.65 | 3.71 | 3.28 | 3.7 | 3.7 | 3.28 |

*Note*. English word frequencies are Zipf values taken from SUBTLEX-UK (van Heuven et al., 2014). Concreteness scores are based on Brysbaert et al. (2014).

A new set of stimuli for task 2 and 3 were selected from the English-Chinese translation norming database (Wen & van Heuven, 2017b). In total, 36 English target words were selected, and these were preceded by three types of two-character Chinese primes: 1) a Chinese translation equivalent, 2) a Chinese word whose first character is the same as the Chinese translation of the English word (partially hidden repetition), 3) a Chinese control word matched in the number of strokes and word frequency with the other primes. Examples of the stimuli are presented in Table 7. For the lexical decision task (task 2), a set of 36 English pseudowords were selected from the British Lexicon Project (Keuleers et al., 2012). These pseudowords matched the target words in length and the first letter. For the final task, 12 additional translation equivalent pairs were included as fillers in order to have an equal number of Yes and No responses. Importantly, the critical stimuli presented in task 2 and 3 were exactly the same. Three counterbalanced lists were created and participants were randomly assigned to one of the lists in task 2 and presented with the same list in task 3.[3] The order of tasks 2 and 3 was not counterbalanced because conducting the explicit translation

task first could potentially impact the participants' performance in the masked priming lexical decision task.

Table 7 Experimental design, mean lexical characteristics of the Chinese primes, and stimulus examples of Experiment 3 task 2 & 3

| | Prime | | | Target |
| --- | --- | --- | --- | --- |
| | *Translation Equivalent* | *Partially Hidden Repetition* | *Control* | |
| Word | 事实 | 事业 | 蜡烛 | fact |
| English Translation | fact | career | candle | |
| Chinese Frequency | 4.58 | 4.49 | 4.50 | |
| Strokes | 15.39 | 15.31 | 15.14 | |

*Note*. Chinese frequencies are Zipf values calculated using the word frequencies from SUBTLEX-CH (Cai & Brysbaert, 2010).

*Procedure*

The procedure of the first task was identical to Experiment 1 of Zhang et al. (2011). Furthermore, the exact same DMDX files from Zhang et al. (2011) were used, except that the timings (in terms of number of frames) had to be modified because the monitor was different from the original monitor used by Zhang et al. (2011). For task 2 a similar procedure as in Experiment 1 was used, except that primes were two-character Chinese words. For the translation recognition task (task 3), each trial began with a fixation (+) presented for 500 ms, followed by a blank screen for 200 ms. Next, a two-character Chinese word was presented for 500 ms followed by a blank screen for 200 ms (Guo et al., 2012). Immediately following this blank screen, the English target word was presented in uppercase. Participants were asked to decide whether the English word was the correct translation of the Chinese word by pressing buttons on a response box (Cedrus) attached to the computer. Participants were instructed to

perform the task as quickly and as accurately as possible. The uppercase string remained on the screen up to a maximum of 3000 ms or until the participants responded. The inter-trial interval was 1000 ms.

**Results**

Table 8 and 9 present the mean reaction times and error rates for each task in Experiment 3. Incorrect responses for the critical stimuli were discarded (task 1: 6.0%; task 2: 3.2%; task 3: 4.1%; error rates for fillers: task 1: 16.4% ; task 2: 9.5%; task 3: 2.0%). Reaction times less than 300 ms or more than 2,500 ms were discarded (task 1: 0.51%; task 2: 0.54%; task 3: 0%). Data of task 1 was analysed using a 2 (control vs. hidden repetition) × 2 (first character vs. second character) repeated measures ANOVA. For task 2, a one-way 3 (translation equivalent vs. partially hidden repetition vs. control) repeated measures ANOVA was conducted, and further comparisons were conducted when there was a significant effect of prime condition. The results of task 3 were analysed with paired t-tests that compared translation equivalent word pairs with translation non-equivalent word pairs and critically comparing the two translation non-equivalent conditions (partially hidden repetition word pairs vs. control word pairs).

*Task 1: Masked hidden repetition priming with English primes and targets*

The analysis of the reaction times revealed no significant effect of hidden repetition priming, $F_1(1, 51) = 1.365$, $p = .248$, $F_2(1, 39) < 1$, $p > .25$, or morpheme position, $Fs < 1$, $ps > .25$, and no significant interaction, $Fs < 1$, $ps > .25$. Also, the error analyses revealed no main effects, $Fs < 1$, $ps > .25$, and no significant interaction, $F_1(1, 51) = 2.72$, $p = .105$, $F_2(1, 39) = 3.03$, $p = .090$. For all analyses, BFs < 0.3.

Table 8 Mean reaction times and error rates (with SD in brackets) in Experiment 3 task 1

| | First Character | | | Second Character | |
|---|---|---|---|---|---|
| | *Repetition* | *Control* | | *Repetition* | *Control* |
| RT (ms) | 944 (197) | 941 (184) | | 949 (202) | 926 (180) |
| Error Rate (%) | 6.73 (8.6) | 5.19 (7.3) | | 5.24 (7.5) | 7.00 (9.2) |

*Task 2: Masked cross-language priming with Chinese primes and English targets*

The RT analysis revealed significant priming effects, $F_1$ (2, 102) = 76.23, $p < .001$, $\eta_p^2 = .599$, BF > 100, $F_2$ (2, 70) = 83.85, $p < .001$, $\eta_p^2 = .706$, BF > 100. Planned comparisons revealed a significant 232 ms translation priming effect, $t_1$ (51) = 9.74, $p < .001$, BF > 100, $t_2$ (35) = 13.44, $p < .001$, BF > 100, and a 44 ms partially hidden repetition priming effect that was significant in the by-subject analysis, $t_1$ (51) = 2.48, $p = .017$, BF = 2.40 and marginally significant in the by-item analysis, $t_2$ (35) = 1.68, $p = .102$, BF = 0.64.

A similar pattern was found in the error analysis. A reliable priming effect was found, $F_1$ (2,102) = 35.63, $p < .001$, $\eta_p^2 = .411$, BF > 100, $F_2$ (2, 70) = 25.56, $p < .001$, $\eta_p^2 = .422$, BF > 100. Planned comparisons revealed that participants made less errors when first presented with translation equivalent primes, $t_1$ (51) = 7.20, $p < .001$, BF > 100, $t_2$ (35) = 6.13, $p < .001$, BF > 100. Participants made also less errors when first presented with partially hidden repetition primes compared to control primes, $t_1$ (51) = 2.78, $p = .008$, BF = 4.74, $t_2$ (35) = 2.86, $p = .007$, BF = 5.73.

Table 9 Mean reaction times and error rates (with SD in brackets) for tasks 2 and 3 of
Experiment 3

| | | *Translation Equivalent* | *Partially Hidden Repetition* | *Control* |
|---|---|---|---|---|
| Task 2 | RT (ms) | 828 | 1016 | 1060 |
| (Masked priming) | | (174) | (182) | (196) |
| | Error Rate (%) | 2.7 | 12.5 | 17.9 |
| | | (4.9) | (9.7) | (14.7) |
| Task 3 | RT (ms) | 541 | 815 | 730 |
| (Translation recognition) | | (106) | (179) | (159) |
| | Error Rate (%) | 1.9 | 8.0 | 2.4 |
| | | (3.5) | (9.6) | (5.0) |

*Task 3: Chinese-English translation recognition*

As expected, participants responded faster to translation equivalent word pairs, which
required a Yes response, than to word pairs that were not translation equivalent and required a
No response, $t_1$ (51) = 11.77, $p < .001$, BF > 100, $t_2$ (35) = 13.84, $p < .001$, BF > 100.
Importantly, the crucial comparison between word pairs that required a No response revealed
a significant interference effect for the partially hidden repetition word pairs compared to
control word pairs, $t_1$ (51) = 6.22, $p < .001$, BF > 100, $t_2$ (35) = 3.44, $p = .002$, BF = 21.62 >
3.

Analyses of error rates revealed no significant differences between translation equivalent and
non-translation equivalent word pairs, $p$s > .25, BFs < 0.3. In line with the response time
analysis, participants made more rejection errors for word pairs with a partially hidden

repetition than to control word pairs, $t_1$ (51) = 5.03, $p$ < .001, BF > 100, $t_2$ (35) = 3.16, $p$ = .003, BF = 11.13> 3.

*Evaluation of the replication attempts of Zhang et al. (2011)*

To evaluate our two replications of Zhang et al. (2011), effect sizes (Cohen's d) and the fixed-effect meta-analysis Bayes factors were calculated for Experiment 2 (Hidden One-character Prime condition with the prime duration of 59 ms), and for the first task of Experiment 3 (the First Character condition in task 1). If the effect size of the original study is outside the 95% confidence interval of these replications, the replication studies fail to replicate the original study (Open Science Collaboration, 2015). A fixed-effects meta-analysis Bayes factor test was also conducted. This test aims to reveal whether an experimental effect is present (Meta BF > 3) or absent (Meta BF < 0.3) on the basis of the original study and the replications (Verhagen & Wagenmakers, 2014). As Table 10 shows, the effect size of Zhang et al. (2011) is clearly outside the 95% confidence intervals of two replications and Meta BFs provide evidence for an absence of a priming effect.

Table 10 Effect sizes and Meta BF of two replications against Zhang et al. (2011)

| Experiment | Effect size d [95% CIs] | Meta BF |
| --- | --- | --- |
| Original study Zhang et al. (2011) | 0.349  [0.53 0.17] | |
| Replication 1: Experiment 2 | - 0.085 [- 0.29 0.12] | 0.019 |
| Replication 2: Experiment 3 | - 0.012 [- 0.17 0.14] | 0.015 |

**GENERAL DISCUSSION**

The original aim of this study was to determine whether the hidden repetition priming effects reported by Zhang et al. (2011) were due to phonological repetition. Surprisingly, across three

experiments we failed to find any evidence of hidden repetition priming effects in the masked priming lexical decision task. Importantly, even the exact replication, with a larger sample size (N = 52, first task of Experiment 3) of the original study by Zhang et al., provided no evidence for a hidden repetition priming effect. The failure to replicate the findings of Zhang et al. (2011) was not due to a lack of statistical power, and furthermore, Bayesian analyses strongly supported the null hypothesis. The fixed-effect meta-analysis Bayes factor tests also provided evidence for no hidden repetition priming. Overall, the experiments failed to replicate the hidden repetition priming effects with English prime and target words in the masked priming lexical decision task.[4]

Although it is not uncommon to fail to replicate an experimental effect (Open Science Collaboration, 2015), it is important to evaluate potential reasons why the present experiments failed to find hidden repetition priming effects. One possible explanation is that participants in the present study did not process the masked primes up to the semantic level. This possibility is very unlikely because the results of the second task of Experiment 3 revealed a significant masked translation priming effect, which is assumed to originate from semantic overlap (Finkbeiner et al., 2004; Schoonbaert et al., 2009).[5]

A second possible explanation is that participants in the present study did not know the Chinese translations of the English words. We would like to emphasise that we followed the same steps as outlined by Wu and Thierry (2010) to validate word translations. Using the "first-translation" method with a group of 28 participants, we created a large-scale Chinese translation database for 1429 English words (Wen & van Heuven, 2017b). The stimuli used in the present study were selected from this database or taken from Zhang et al. (2011). The accuracy rate in the study by Wen and van Heuven (2017b) was 86.7%. Therefore, we are confident that our participants knew the Chinese translations of the English words. In addition, we observed partially hidden repetition priming effects and translation priming

effects in task 2 of Experiment 3, which further indicates that our participants knew the Chinese translations.

A third possible explanation is that the linguistic profile of the bilingual participants (e.g., English proficiency, immersion experience) in the present experiments was different from that of the bilingual participants in Zhang et al. (2011) even though they were recruited from the same university. In terms of English proficiency, the average self-rated proficiency scores in Zhang et al. (2011) did not different from the participants who participated in the three experiments of the present study (all $ps > .25$), ruling out differences in proficiency as a possible explanation. However, the impact of English proficiency cannot be completely ruled out because self-assessments of language proficiency may be not as reliable as objective proficiency measures (Khare, Verma, Kar, Srinivasan, & Brysbaert, 2013; Lemhöfer & Broersma, 2012; Wen & van Heuven, 2017b). Although the English language proficiency of the bilinguals in the present study were objectively measured using an English vocabulary test (LexTALE, Lemhöfer & Broersma, 2012), this test was not used in Zhang et al. (2011) and therefore it is not possible to compare the objective English proficiency measures across the studies. Nevertheless, the participants' LexTALE scores did not correlate with the hidden repetition priming effects in our exact replication (task 1 of Experiment 3), r = 0.02, $p$ =.90. Thus, it is unlikely that English proficiency differences can explain our failure to find hidden repetition priming effects. To explore the role of immersion, the immersion experience (number of months in the UK) was entered as a fixed factor in the mixed-effect modelling analysis of Zhang et al. (2011), and task 1 of Experiment 3. The results revealed that immersion experience is not a significant predictor in both analyses (i.e., Zhang et al.: $t =$ 1.131, $p > .26$; task 1 of Experiment 3: $t$ <1). Therefore, immersion experience can also not explain our failure to replicate. Another possibility is that the participants in the different studies were exposed to different teaching English methods. However, we are not aware of

the recent fundamental changes in teaching English in China. English language education in China is still dominated by traditional teaching methods (e.g., grammar-translation method and audiolingualism approach), even though new methods (e.g., communicative language teaching) have been introduced (Hu, 2001, 2002a, 2002b; Yu, 2001). Taken together, the bilinguals in Zhang et al. (2011) and the present study had very similar linguistic profiles as assessed by language questionnaires and the same English immersion experience. Thus, it is unlikely that participants' linguistic profile can explain the contradictory findings of the present study and Zhang et al. (2011).

A fourth possible explanation is that there were errors in the original analyses. To explore this possibility, we reanalysed the raw data of Zhang et al. (2011). No errors were found in the original analyses. Furthermore, a re-analysis using mixed-effects modelling also revealed similar findings as presented in Zhang et al. (2011). Crucially, the Bayesian factors also supported the presence of priming effects, BF > 3. Close inspection of the means for the different conditions of Zhang et al. indicated that reaction times for the control condition of the hidden first character (e.g., wall-thing) were significantly slower than those for the control condition of the hidden second character (e.g., town-thing), 1005 ms vs. 946 ms, $t_1 (30) = 2.337$, $p = .026$, $t_2 (30) = 2.118$, $p = .041$. Because the same set of target words were used and both groups of control words were matched in terms of word frequency and length, the differences in reaction times between the two controls might be a false positive. Thus, our re-analysis confirmed the hidden repetition priming effects in the original study.

The final possible explanation is that the effect size of Zhang et al. (2011) was small and that therefore the power was not very high, which would lead to a low chance of reproducing the effects. This seems to be the most plausible account because despite a strong priming effect (i.e., 75 ms), the standardized effect size was indeed small (i.e., d = 0.349). Based on this small effect size, a power analysis indicated that Zhang's et al (2011) experiment had only a

power of 0.61 (N = 32, one-tailed paired t-test). Because power indicates the probability of successfully rejecting the null hypothesis (i.e., 1 minus type II error), the chance of finding significant results in three successive replication experiments would be $0.61^3 = 0.227$ at best (Ioannidis & Trikalinos, 2007). However, this probability of successful replication might actually be an over-estimation because the observed power is an overestimate of the true power (Vasishth & Nicenboim, 2016). This is because the true effect size of hidden repetition priming effects may be even smaller than what was observed in Zhang et al. (2011). In fact, the effect size of partially hidden repetition priming effects (d = 0.23) observed with masked Chinese primes in the second task of Experiment 3 was smaller than the effect size of hidden repetition priming effects (masked English primes) in Zhang et al. (2011). This is remarkable because why would hidden repetition priming effects involving prime and target translations be larger than *partially* hidden repetition priming effects, which only requires targets to be translated. Furthermore, a recent meta-analysis reported that the mean effect size of L2-L1 masked translation priming studies was 0.31 (Wen & van Heuven, 2017a). Because this translation priming effect is due to semantic overlap (Finkbeiner et al., 2004; Schoonbaert et al., 2009), it is unlikely that a hidden repetition priming effect, which can only occur because of form repetition through translation, would lead to a larger effect size than translation priming. Taken together, we failed to replicate the hidden repetition priming effects because the effect size of the original study was small and its power was low.

To confirm this conclusion, simulations were conducted to estimate the replicability of hidden repetition priming effects. Vasishth and Nicenboim (2016) provided an R function (sim_data.R) that can simulate reaction time data of experiments involving a Latin-square design. Using the relevant parameters from the results of Zhang et al. (2011) (i.e., standard deviations in by-item and by-subject analysis, the correlations of reaction times in two experimental conditions, and the intercept of the mixed-effects modelling analysis), 1,000

experiments were simulated with 32 participants, 10 items per cell for a true priming effect (in millisecond instead of Cohen's d) estimated at 4 ms, 8 ms, and 12 ms respectively. These estimates were chosen because the average L2-L1 translation priming effect was 11 ms based on all translation priming experiments used in a recent meta-analysis (Wen & van Heuven, 2017a). Therefore, it is unlikely that the true effect of the hidden repetition priming is larger than 12 ms.

The analysis of the simulations are presented in Figure 1 with each point representing a significant effect and the dashed line depicting the true effect (e.g., 4 ms, 8 ms, and 12 ms). Figure 1 also shows the proportions of significant effects (numbers in black for facilitatory effects and numbers in grey for inhibitory effects). For example, if the true priming effect was 8 ms, 51% of the 1,000 simulated experiments had a significant facilitatory effect and 5% of the simulated experiments showed a significant inhibitory effect. Therefore, the true power with the true priming effect of 8 ms is estimated to be 0.51. In other words, if the true effect is as small as 8 ms, the chance of finding significant results in a replication experiment with 32 participants and 10 items per cell would be 51%. Although a limitation of this simulation is that the magnitude of the true effect is unknown, the simulations provide a general picture of the relationship between effect size, power and replications. Critically, an important observation of the results of these simulations is that none of the simulated experiments observed a priming effect as large as 75 ms found in Zhang et al. (2011). Therefore, these simulations confirm that the priming effect observed in Zhang et al. (2011) is an over-estimation of the true effect.
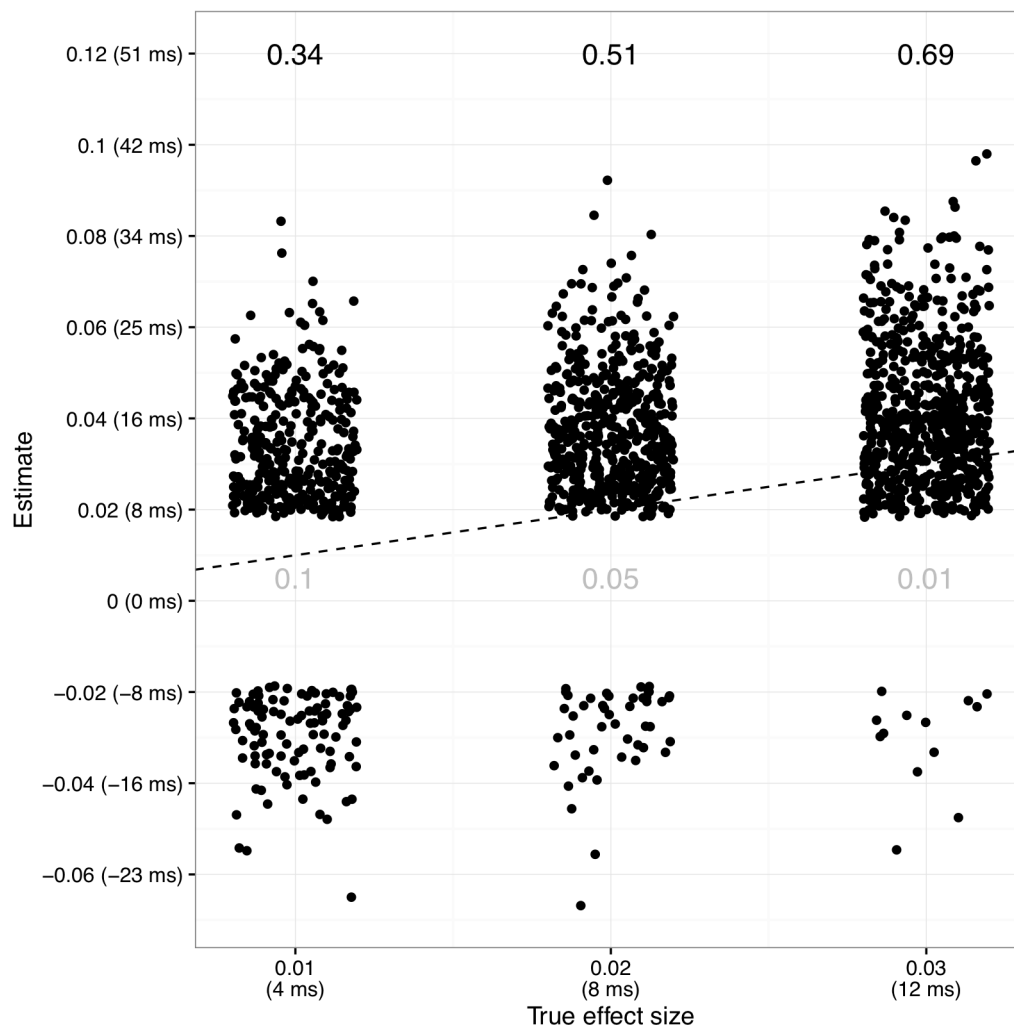
Figure 1 Results of simulated experiments with a true priming effect of 4, 8 12 ms; each point representing a significant effect and the dashed line depicting the true effect (e.g., 4 ms, 8 ms, and 12 ms); numbers in black are the proportions of significant facilitatory effects numbers in grey are the proportions of significant inhibitory effects.

Taken together, our systematic replication experiments did not provide behavioural evidence for hidden repetition priming effects in the masked priming lexical decision task. The failure to find evidence for the activation of the Chinese translations of masked English primes could be a limitation of behavioural measures, which might not be sensitive enough to detect the hidden repetition priming effects. Electrophysiological measures might be more sensitive

than behavioural measures because electrophysiological data can unravel the time course of bilingual word processing, which cannot be captured by behavioural data (Casaponsa, Carreiras, & Duñabeitia, 2015; Duñabeitia, Dimitropoulou, Dowens, Molinaro, & Martin, 2016; van Hell & Kroll, 2013). However, as discussed in the Introduction, ERP studies have suggested that masked L2 primes do not activate L1 translation forms (e.g., Hoshino et al., 2010; Midgley et al., 2009; Schoonbaert et al., 2011). Therefore, it seems unlikely that electrophysiological measures would be able to reveal hidden repetition priming effects in the masked lexical decision task.

Although the present study failed to replicate the hidden repetition priming effects observed in reaction times (Zhang et al., 2011), it is important to point out that this conclusion does not contradict the electrophysiological findings by Wu and Thierry (Thierry & Wu, 2007; Wu & Thierry, 2010, 2012a) and our own electrophysiological findings (Wen, Filik, & van Heuven, in press). An important difference is that these electrophysiological studies used the semantic relatedness judgment task, whereas the behavioural experiments presented here and in Zhang et al. (2011) used the lexical decision task. These two experimental tasks differ in an number of aspects, such as the task demands and importantly the visibility of the prime (i.e., both the prime word and target word were visible in Wu and Thierry, whereas only the target word was visible in the present study). Thus, prime durations (i.e., 500 ms in Wu and Thierry vs. 59 ms in the present study) differ as well as the ISI (i.e., 400 or 600 ms on average in Wu and Thierry vs. 24 ms in the present study). It is therefore very difficult to directly compare these studies, so our key focus in this study is on the comparison between the present study and Zhang et al. (2011) because they both involved the same task and procedures.

To conclude, the present study suggests that there is no behavioural evidence that the Chinese translations of the English primes and targets are *both* activated in the masked priming lexical decision task with unbalanced Chinese-English bilinguals. However, the partially hidden

repetition priming effect observed in task 2 of experiment 3 suggests that the Chinese translation of the English target word is activated. [6] The partially hidden repetition priming effects were significant in the by-subject analysis and marginally significant in the by-item analysis. We also conducted mixed-effects modelling which revealed, in contrast to the ANOVA results, a significant priming effect ($t = 3.192$, $p = .002$). However, the effect size of the partially hidden repetition priming effects was small ($d = 0.23$), so a critical question is whether the partial hidden repetition priming effects observed in the present study is replicable. Using a different set of items with an identical procedure, Zhang (2013) also reported partially hidden repetition priming effects when masked Chinese primes shared a character with the Chinese translations of English target words. This priming effect had a similar effect size ($d = 0.29$). Thus, these data provide additional support for partially hidden repetition priming effects. Further support for the activation of the Chinese translations of English target words comes from the results of the third task of Experiment 3. In this translation recognition task with unmasked Chinese-English word pairs, Chinese character repetitions lead to an interference effect ($d = -0.49$). This can only be explained by assuming that the Chinese translations of English target words were activated.

To understand the mechanisms that could account for the partially hidden repetition priming effects, it is useful to look at some of the potential representations that are activated during the masked priming task. In line with the model proposed by Voga and Grainger (2007), we assume that there is no inhibition between different-script languages. Figure 2 presents a model of potential representations that are activated just after the target word has been presented for the repetition and control conditions in task 2 of Experiment 3. A partially hidden repetition priming effect could in this case occur when a Chinese prime activates Chinese word neighbours (words with shared characters). For example, the Chinese prime 事业 [career] activates the Chinese word 事实 [fact] via the shared character 事. These

activated Chinese words (e.g., 事实 [fact]) could subsequently activate their semantic representations so that the recognition of the English translation equivalent target word (e.g., fact) is facilitated through feedback from semantics to the English word forms. However, a potential issue is that Chinese words generally share one of their characters with many other words (on average 67 first-character neighbours for the primes in our study). For example, the Chinese character 事 occurs in the words 事情 [thing], 事件 [incident],事故 [accident], 事迹 [deeds], 事物 [object], 事务 [routines], 故事 [story],公事 [business] ,私事 [personal issue]. Assuming competition at the lexical level (lateral inhibition), the prime (e.g., 事业 [career]) would suppress these other words (see Figure 2A). To explore the impact of the number of prime neighbours, we entered the number of prime neighbours as a fixed factor in the mixed effect modelling analysis of task 2 in Experiment 3. If a Chinese prime activates Chinese word neighbours up to the semantic level, we expected that the number of prime neighbours influenced the partially hidden repetition priming effects. However, the results revealed that the number of prime neighbours was not a significant predictor ($t < 1$). Thus, it is seems unlikely that same-character neighbours are strongly activated and that the Chinese translation of the target (e.g., 事实 [fact]) would activate its semantic features sufficiently so that English target word recognition would be facilitated.

Another possible explanation for partially hidden repetition priming is that English target words activate their Chinese word form translations. When the translation of the target (e.g., 事实 [fact]) is activated, its activation receives an extra boost because it shares the first character with the prime (e.g., 事业 [career], see Figure 2A. Importantly, such a boost would not happen in the control condition (see Figure 2B), and it would not happen in the hidden repetition priming if the Chinese translations of English primes are not activated. Thus, in partially hidden repetition priming, the activation of the shared character increases the

activation level of the Chinese translation of the target word and its semantic features, which could then in turn facilitate the target word recognition through feedback from the semantic level to the English word form level.
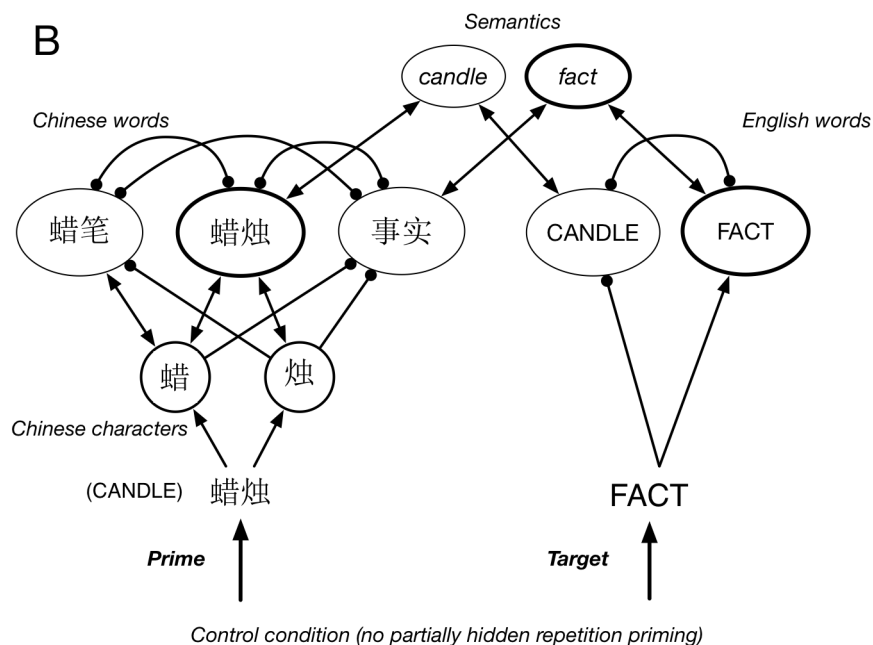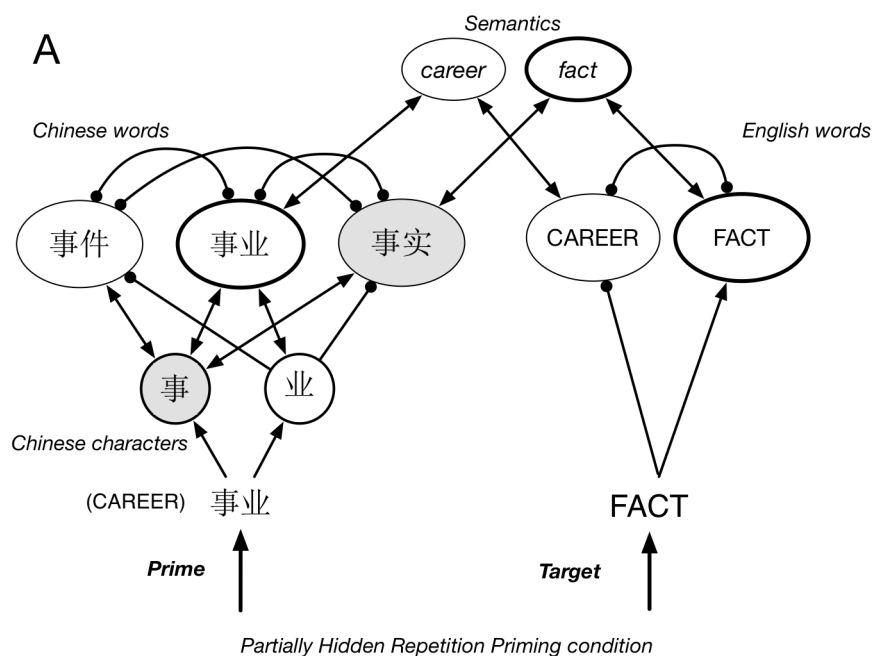


Figure 2 Model illustrating a hypothetical pattern of activated representations and connections for A) partially hidden repetition priming effect, and B) the control condition in the masked priming lexical decision task. Arrows with triangle heads indicate excitation and arrows with circular heads indicate inhibition. Grey circles indicate the crucial representations that play a role in partially hidden repetition priming.

It is unclear which of the above mechanisms is able to account for partially hidden repetition priming effects. To help clarifying the partial repetition priming effects as well as the other findings of our experiments, we implemented a connectionist model of visual word recognition in proficient Chinese-English bilinguals based on the ideas presented in Figure 2. This Chinese-English Interactive Activation Model (CE-IAM) uses the same activation functions as in the Interactive Activation (IA) model (McClelland & Rumelhart, 1981; McClelland & Rumelhart, 1988) and the bilingual interactive activation model (Dijkstra & van Heuven, 2002; van Heuven et al., 1998). The model has a Chinese visual word recognition component and an English visual word recognition component. Chinese and English words are bidirectionally connected to semantic nodes to allow the model to translate between Chinese and English (see Figure 3). Our goal was to keep the model as simple as possible in order to investigate if such a simple model is able to account for our masked priming results. In particular, our focus was on the activation of word form translations and therefore the model only contained orthographic representations and simple semantic representations. A letter layer was not included in the model to keep the model as simple as possible. Therefore, the English word recognition component of the model included only a layer of English word forms. The Chinese part contained a layer of Chinese word forms and a layer with sets of first position and second position Chinese characters in order for the model to recognise two-character Chinese words. The semantic layer linked the Chinese and English word forms through a single semantic node representing the meaning of the words. The Chinese and English lexicons, as well as the translations used to create semantic nodes and to connect semantic nodes to word forms, were taken from the English-Chinese translation norming database (Wen & van Heuven, 2017b). The parameters of the model were based on the IA model and set to values that created a stable activation pattern. Furthermore, they were set so that a Chinese word was recognised as fast as an English word with the same word

frequency (see Appendix A for model details and an overview of the parameter settings).

Thus, the model simulated a proficient balanced bilingual conducting an English lexical

decision task. Parameters were kept the same across all simulations unless reported

otherwise. A fixed word recognition threshold of 0.70 at the English word form layer was

used in all simulations and primes were presented for 2 cycles to simulate masked priming

(similar to simulations with the IA model, e.g., van Heuven, Dijkstra, Grainger, & Schriefers,

2001).

 The model assumes that different script languages are only interconnected through the

semantic layer (see also Voga & Grainger, 2007). However, CE-IAM is still fully interactive

and the different script languages interact through the semantic layer and therefore the model

is consistent with the assumptions of other bilingual models of visual word recognition

(Dijkstra & van Heuven, 2002). Simulations showed that word form translations in the other

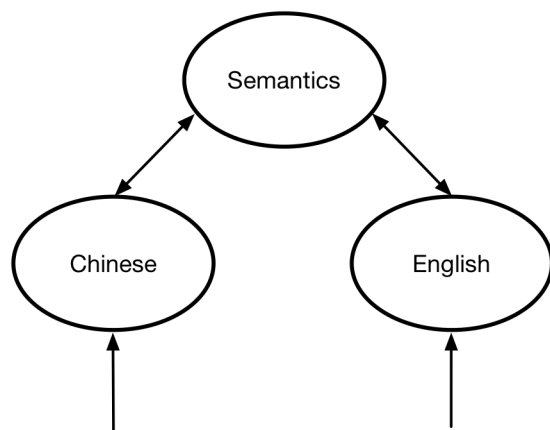language were quickly activated in the model.



Figure 3 Chinese-English Interactive Activation Model (CE-IAM)

We first conducted simulations of Experiment 3 (task 2) by presenting all prime-target pairs

to the model. The results revealed significant translation priming effects (1.44 cycles, $t$ (35) =

17.20, $p < .001$, see Figure 4A) and importantly, partially hidden repetition priming effects

(0.58 cycles, $t$ (35) = 7.00, $p < .001$). Interestingly, simulations of Experiment 2 did not reveal any hidden repetition priming effect (identical response times for hidden repetition prime and control conditions).[7] Thus, these simulation results are compatible with the all of our experimental findings. However, our model simulated a balanced bilingual, whereas our experiments were conducted with unbalanced proficient bilinguals. It is unclear yet, what the best way is to simulate unbalanced bilinguals. Therefore, we conducted simulations with two potential model variants of unbalanced bilinguals. In one model variant we reduced the connection weights (excitation) between English words and semantics (both forward and backward) and in another model we only reduced all English word frequencies. Simulations with both models showed, however, a similar pattern of results as reported above and presented in Figure 4. Thus, our simulations revealed that English proficiency does not seem to impact partially hidden and hidden repetition priming effects.

To investigate whether translation form activation is crucial for partially hidden priming effects, we conducted simulations with the same stimuli (task 2 of Experiment 3) in a model without feedback from semantics to Chinese word forms. The model still contained excitatory connections from Chinese word forms to semantics. Thus, if simulations with this version of the model reveal partially hidden repetition priming effects, the source must be pre-activation of semantic representations. However, if no partially hidden repetition priming effects are observed, it would suggest that target word translation must be the source of the partially hidden repetition priming (as this requires a feedback connection from semantics to Chinese word forms). The simulations revealed again significant translation priming effects (1.28 cycles, $t$ (35) = 16.88, $p < .001$, see Figure 4B). Critically, the simulations showed no partially hidden repetition priming effects at all (identical response times for the partially hidden repetition and control conditions). Thus, according to the model, the excitation of Chinese translation forms (through the connection from semantics to Chinese words) is

essential for partially hidden repetition priming effects. Future simulation work will explore the ability of the model to account for other masked priming effects reported in the literature.

A: Simulation Results (Semantic nodes to Chinese words excitation)
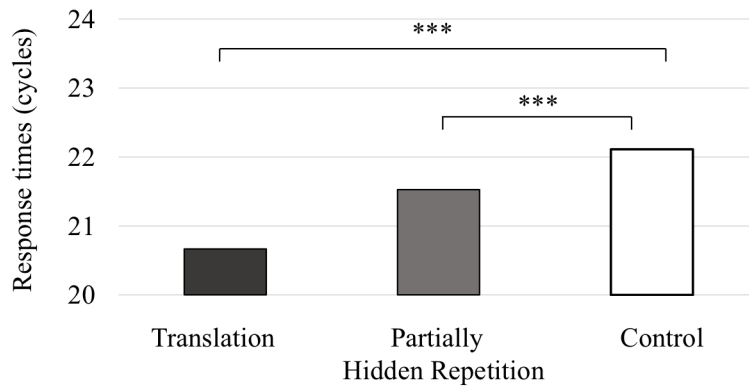
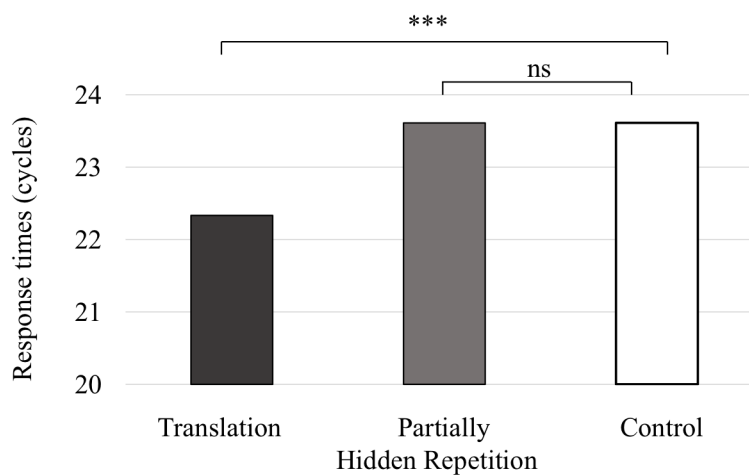B: Simulation Results (no Semantic nodes to Chinese words excitation)

Figure 4 Results of simulations with CE-IAM (A: Simulations with feedback from semantics to Chinese word forms; B: Simulations without feedback from the semantics to Chinese word forms). *** $p < .001$

To summarize, the three behavioural experiments in the present study with in total 118 Chinese-English bilinguals provided no evidence for hidden repetition priming in the masked priming lexical decision task with unbalanced Chinese-English bilinguals. However, a classic

masked translation priming effect was found, indicating that the Chinese primes were processed up to the semantic level. Furthermore, partially hidden repetition priming effects were obtained with masked and unmasked Chinese primes, suggesting that target words activate their translations. Our computational model was able to simulate partially hidden repetition priming effects and the absence of hidden repetition priming effects. Importantly, simulations showed that partially hidden repetition priming effects in the model occurred through feedback from semantics to Chinese words, which provides further support that the activation of word form translations is limited to target words in the masked priming paradigm.

**ACKNOWLEDGMENTS**

**FOOTNOTES**

1. After completing the English lexical decision task, participants also conducted a Chinese lexical decision task that included Chinese targets with masked Chinese primes. The Chinese words were translation equivalents of the English words. However, because the majority of participants realised that the Chinese words were translation equivalents of the English words, the Chinese data should be interpreted with caution. For the phonological repetition condition, all error rates and reaction times analyses revealed no significant differences, all $ts < 1$, $ps > .25$, all BFs $< 0.3$. For the orthographic and phonological repetition prime, significant differences were found in the reaction times in the by-subject analysis, $t_1 (31) = 3.000$, $p = .005$, BF $>$ 3; $t_2 (39) = 1.558$, $p = .127$, BF $= 0.517$. Similarly, error rates also showed significant effects in the by-subject analysis, $t_1 (31) = 2.048$, $p = .049$, BF $= 1.184$; $t_2 (39) = 1.621$ $p = .113$, BF $= 0.566$.

2. To avoid excluding a larger percentage of the data, we used 3000 ms instead of 2500 ms as the outlier criteria. However, the pattern remained the same when the outlier criteria was set at 2500 ms.

3. For task 2 and 3, list 1 had 18 participants, whereas lists 2 and 3 had 17 participants. The results did not change when we excluded participant 18 from list 1.

4. We decided to present the ANOVA analyses because the BF is based on the results of the ANOVA analyses. Reaction times were also analysed using mixed-effects modelling in all three experiments, which revealed the same findings.

5.  As suggested by one of the reviewers, a related issue is whether participants varied in prime awareness, which might lead to facilitatory and inhibitory effects across participants and therefore cancelling the priming effects. We have asked participants in all the three experiments whether they were aware of the masked English primes after they conducted the experiment. None of participants reported prime awareness, and they were all surprised. Thus, it seems unlikely that our participants varied in prime awareness. In another experiment using the same procedures in our lab (same computer, same lighting and prime duration, masking) (Experiment 3 in Wen, 2017), we informed participants of the masked English primes and asked them to name the primes after they had conducted a lexical decision task with the same stimuli. Although participants tried their best, correct identification accuracy was only 10.2%.

6.  As suggested by one of the reviewers, the different effects in tasks 1 and 2 of Experiment 3 may originate in bilinguals' different awareness of L1 and L2 primes. As observed by Wang and Forster (2015), Chinese-English bilinguals have higher awareness for L1 words (Chinese) than for L2 words (English). However, in the literature, researchers normally use the same procedures when investigating priming effects from L1 to L2 and from L2 to L1 (see Dimitropoulou, Duñabeitia, & Carreiras, 2011). Therefore, we used the same procedure in tasks 1 and 2 in order to be able to compare the results. Although our procedure may not be perfect, the key focus of task 2 with masked Chinese primes was to investigate whether English targets were translated or not. Therefore, the potential issue of differences in awareness of L1 and L2 primes does not impact our interpretation.

7.    Because the stimuli used in task 1 were taken from Zhang et al. (2011) and not all

these stimuli were included in the English-Chinese translation norming study (Wen &

van Heuven, 2017b), we used the stimuli from Experiment 2 (hidden two-character

prime condition) to simulate hidden repetition priming.

**References**

Allen, D. B., Conklin, K., & van Heuven, W. J. B. (2015). Making sense of the Sense Model: Translation priming with Japanese-English bilinguals. *The Mental Lexicon, 10*(1), 32-52. doi: 10.1075/ml.10.1.02all

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience, 5*(1), 34-44. doi: 10.1162/jocn.1993.5.1.34

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46*(3), 904-911. doi: 10.3758/s13428-013-0403-5

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE, 5*(6), e10729. doi: 10.1371/journal.pone.0010729

Casaponsa, A., Carreiras, M., & Duñabeitia, J. A. (2015). How do bilinguals identify the language of the words they read? *Brain Research, 1624*, 153-166. doi: 10.1016/j.brainres.2015.07.035

Chen, B., Zhou, H., Gao, Y., & Dunlap, S. (2014). Cross-language translation priming asymmetry with Chinese-English bilinguals: A test of the sense model. *Journal of Psycholinguistic Research, 43*(3), 225-240. doi: 10.1007/s10936-013-9249-3

de Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30*(1), 90-123. doi: 10.1016/0749-596X(91)90012-9

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*. doi: 10.3389/fpsyg.2014.00781

Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. De Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 178-201). New York: Oxford University.

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. doi: 10.1017/S1366728902003012

Dimitropoulou, M., Duñabeitia, J. A., & Carreiras, M. (2011). Two words, one meaning: Evidence of automatic co-activation of translation equivalents. *Frontiers in Psychology, 2*. doi: 10.3389/fpsyg.2011.00188

Duñabeitia, J. A., Dimitropoulou, M., Dowens, M. G., Molinaro, N., & Martin, C. (2016). The electrophysiology of the bilingual brain. In R. R. Heredia, J. Altarriba & A. B. Cieślicka (Eds.), *Methods in Bilingual Reading Comprehension Research* (pp. 265-312). New York: Springer.

Duñabeitia, J. A., Dimitropoulou, M., Uribe-Etxebarria, O., Laka, I., & Carreiras, M. (2010). Electrophysiological correlates of the masked translation priming effect with highly proficient simultaneous bilinguals. *Brain Research, 1359*, 142-154. doi: 10.1016/j.brainres.2010.08.066

Ferré, P., Sánchez-Casas, R., & Guasch, M. (2006). Can a horse be a donkey? Semantic and form interference effects in translation recognition in early and late proficient and nonproficient Spanish-Catalan bilinguals. *Language Learning, 56*(4), 571-608. doi: 10.1111/j.1467-9922.2006.00389.x

Finkbeiner, M., Forster, K. I., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language, 51*(1), 1-22. doi: 10.1016/j.jml.2004.01.004

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116-124. doi: 10.3758/BF03195503

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(5), 1122-1139. doi: 10.1037/0278-7393.23.5.1122

Grainger, J., & Frenck-Mestre, C. (1998). Masked priming by translation equivalents in proficient bilinguals. *Language and Cognitive Processes, 13*(6), 601-623. doi: 10.1080/016909698386393

Guasch, M., Sánchez-Casas, R., Ferré, P., & García-Albea, J. E. (2008). Translation performance of beginning, intermediate and proficient Spanish-Catalan bilinguals: Effects of form and semantic relations. *The Mental Lexicon, 3*(3), 289-308. doi: 10.1075/ml.3.3.03gua

Guo, T., Misra, M., Tam, J. W., & Kroll, J. F. . (2012). On the time course of accessing meaning in a second language: An electrophysiological investigation of translation recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(5), 1165-1186. doi: 10.1037/a0028076

Hoshino, N., Midgley, K. J., Holcomb, P. J., & Grainger, J. (2010). An ERP investigation of masked cross-script translation priming. *Brain Research, 1344*, 159-172. doi: 10.1016/j.brainres.2010.05.005

Hu, G. (2001). *English language teaching in the People's Republic of China*: National Institute of Education, Nanyang Technological University.

Hu, G. (2002a). Potential cultural resistance to pedagogical imports: The case of communicative language teaching in China. *Language, Culture and Curriculum, 15*(2), 93-105. doi: 10.1080/07908310208666636

Hu, G. (2002b). Recent important developments in secondary English-language teaching in the People's Republic of China. *Language, Culture and Curriculum, 15*(1), 30-49. doi: 10.1080/07908310208666631

Ioannidis, J. P. A., & Trikalinos, T. A. (2007). An exploratory test for an excess of significant findings. *Clinical Trials, 4*(3), 245-253. doi: 10.1177/1740774507079441

Jiang, N. (1999). Testing processing explanations for the asymmetry in masked cross-language priming. *Bilingualism: Language and Cognition, 2*(1), 59-75.

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British lexicon project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods, 44*(1), 287-304. doi: 10.3758/s13428-011-0118-4

Khare, V., Verma, A., Kar, B., Srinivasan, N., & Brysbaert, M. (2013). Bilingualism and the increased attentional blink effect: Evidence that the difference between bilinguals and monolinguals generalizes to different levels of second language proficiency. *Psychological Research, 77*(6), 728-737. doi: 10.1007/s00426-012-0466-4

Kroll, J. F., & Tokowicz, N. (2001). The development of conceptual representation for words in a second language. In J. L. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 49-71). Malden: Blackwell.

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

Ma, F., Chen, P., Guo, T., & Kroll, J. F. (2017). When late second language learners access the meaning of L2 words: Using ERPs to investigate the role of the L1 translation

equivalent. *Journal of Neurolinguistics, 41*, 50-69. doi:
10.1016/j.jneuroling.2016.09.006

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context
effects in letter perception, Part 1: An account of basic findings. *Psychological
Review, 88*(5), 375-405. doi: 10.1037/0033-295X.88.5.375

McClelland, J. L., & Rumelhart, D. E. (1988). *Explorations in Parallel Distributed
Processing: A Handbook of Models, Programs, and Exercises*. Cambridge, MA: MIT
press.

Meade, G., Midgley, K. J., Sevcikova Sehyr, Z., Holcomb, P. J., & Emmorey, K. (2017).
Implicit co-activation of American Sign Language in deaf readers: An ERP study.
*Brain and Language, 170*, 50-61. doi: 10.1016/j.bandl.2017.03.004

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009). Masked repetition and translation
priming in second language learners: A window on the time-course of form and
meaning activation using ERPs. *Psychophysiology, 46*(3), 551-565. doi:
10.1111/j.1469-8986.2009.00784.x

Moldovan, C. D., Demestre, J., Ferré, P., & Sánchez-Casas, R. (2016). The role of meaning
and form similarity in translation recognition in highly proficient balanced bilinguals:
A behavioral and ERP study. *Journal of Neurolinguistics, 37*, 1-11. doi:
10.1016/j.jneuroling.2015.07.002

Moldovan, C. D., Sánchez-Casas, R., Demestre, J., & Ferré, P. (2012). Interference effects as
a function of semantic similarity in the translation recognition task in bilinguals of
Catalan and Spanish. *Psicologica: International Journal of Methodology and
Experimental Psychology, 33*(1), 77-110.

Morey, R. D., & Rouder, J. N. (2016). BayesFactor 0.9.12-2 Comprehensive R Archive
Network. from https://cran.r-project.org/web/packages/BayesFactor/index.html

Morford, J. P., Wilkinson, E., Villwock, A., Piñar, P., & Kroll, J. F. (2011). When deaf signers read English: Do written words activate their sign translations? *Cognition, 118*(2), 286-292. doi: 10.1016/j.cognition.2010.11.006

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716. doi: 10.1126/science.aac4716

Schoonbaert, S., Duyck, W., Brysbaert, M., & Hartsuiker, R. J. (2009). Semantic and translation priming from a first language to a second and back: Making sense of the findings. *Memory & Cognition, 37*(5), 569-586. doi: 10.3758/MC.37.5.569

Schoonbaert, S., Holcomb, P. J., Grainger, J., & Hartsuiker, R. J. (2011). Testing asymmetries in noncognate translation priming: Evidence from RTs and ERPs. *Psychophysiology, 48*(1), 74-81. doi: 10.1111/j.1469-8986.2010.01048.x

Thierry, G., & Wu, Y. J. (2004). Electrophysiological evidence for language interference in late bilinguals. *Neuroreport, 15*(10), 1555-1558. doi: 10.1097/01.wnr.0000134214.57469.c2

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign language comprehension. *Proceedings of the National Academy of Sciences of the USA, 104*(30), 12530-12535. doi: 10.1073/pnas.0609927104

van Hell, J. G., & Kroll, J. F. (2013). Using electrophysiological measures to track the mapping of words to concepts in the bilingual brain: A focus on translation. In J. Altarriba & L. Isurin (Eds.), *Memory, Language, and Bilingualism: Theoretical and Applied Approaches* (pp. 126-160). New York: Cambridge University Press.

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. doi: 10.1006/jmla.1998.2584

van Heuven, W. J. B., Dijkstra, T., Grainger, J., & Schriefers, H. (2001). Shared neighborhood effects in masked orthographic priming. *Psychonomic Bulletin & Review, 8*(1), 96-101. doi: 10.3758/bf03196144

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

Vasishth, S., & Nicenboim, B. (2016). Statistical methods for linguistic research: Foundational ideas-Part I. *Language and Linguistics Compass, 10*(8), 349-369. doi: 10.1111/lnc3.12201

Verhagen, J., & Wagenmakers, E.-J. (2014). Bayesian tests to quantify the result of a replication attempt. *Journal of Experimental Psychology: General, 143*(4), 1457-1475. doi: 10.1037/a0036731

Voga, M., & Grainger, J. (2007). Cognate status and cross-script translation priming. *Memory & Cognition 35*(5), 938-952. doi: 10.3758/BF03193467

Wang, X., & Forster, K. I. (2010). Masked translation priming with semantic categorization: Testing the Sense Model. *Bilingualism: Language and Cognition, 13*(3), 327-340. doi: 10.1017/S1366728909990502

Wang, X., & Forster, K. I. (2015). Is translation priming asymmetry due to partial awareness of the prime? *Bilingualism: Language and Cognition, 18*(4), 651-669. doi: 10.1017/S1366728914000650

Wen, Y. (2017). *Behavioural and electrophysiological investigations of Chinese translation activation during English word recognition in Chinese-English bilinguals.* (Unpublished Ph.D. Thesis), University of Nottingham.

Wen, Y., Filik, R., & van Heuven, W. J. B. (in press). Electrophysiological dynamics of Chinese phonology during visual word recognition in Chinese-English bilinguals. *Scientific Reports*.

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

Wen, Y., & van Heuven, W. J. B. (2017b). Chinese translation norms for 1429 English words. *Behavior Research Methods, 49*(3), 1006-1019. doi: 10.3758/s13428-016-0761-x

Wu, Y. J., Cristino, F., Leek, C., & Thierry, G. (2013). Non-selective lexical access in bilinguals is spontaneous and independent of input monitoring: Evidence from eye tracking. *Cognition, 129*(2), 418-425. doi: 10.1016/j.cognition.2013.08.005

Wu, Y. J., & Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *The Journal of Neuroscience, 30*(22), 7646-7651. doi: 10.1523/JNEUROSCI.1602-10.2010

Wu, Y. J., & Thierry, G. (2012a). How reading in a second language protects your heart. *The Journal of Neuroscience, 32*(19), 6485-6489. doi: 10.1523/JNEUROSCI.6119-11.2012

Wu, Y. J., & Thierry, G. (2012b). Unconscious translation during incidental foreign language processing. *NeuroImage, 59*(4), 3468-3473. doi: 10.1016/j.neuroimage.2011.11.049

Xia, V., & Andrews, S. (2015). Masked translation priming asymmetry in Chinese-English bilinguals: Making sense of the Sense Model. *The Quarterly Journal of Experimental Psychology, 68*(2), 294-325. doi: 10.1080/17470218.2014.944195

Yu, L. (2001). Communicative language teaching in China: Progress and resistance. *TESOL Quarterly, 35*(1), 194-198. doi: 10.2307/3587868

Zhang, T. (2013). *Automatic translation from English to Chinese and subsequent processing of the Chinese translations.* (Unpublished Ph.D. Thesis), University of Nottingham.

Zhang, T., van Heuven, W. J. B., & Conklin, K. (2011). Fast automatic translation and morphological decomposition in Chinese-English bilinguals. *Psychological Science, 22*(10), 1237-1242. doi: 10.1177/0956797611421492

**APPENDIX A**

The Chinese-English Interactive Activation Model (CE-IAM) was implemented in Python. The model has four layers: Chinese characters, Chinese words, Semantic nodes, and English words. The layer of Chinese characters contains two sets of characters, one set for the first character position and one set for the second character position. The model's activation functions are the same as in the interactive activation (IA) model (McClelland & Rumelhart, 1981, 1988). Furthermore, core parameters identical to those the IA model, such as decay (-0.07), fgain (0.05), and the way in which resting-level activations are calculated are also the same as in the IA model. The lexicons consist of 2-character Chinese words and their English translations taken from Wen and van Heuven (2017b). In total, the model contained 1336 English word forms, 1259 Chinese word forms, 1336 semantic nodes, 824 first position Chinese characters, and 741 second position Chinese characters.

Below is an overview of the parameters that determine the weights between the nodes in the model.

| | |
|---|---|
| Excitation Chinese character to Chinese words | 0.14 |
| Inhibition Chinese character to Chinese words | -0.04 |
| Excitation from Chinese words to Chinese characters | 0.30 |
| | |
| Lateral inhibition (English words) | -0.21 |
| Lateral inhibition (Chinese words) | -0.21 |
| Lateral inhibition (Chinese characters) | 0 |
| | |
| Excitation from English words to semantic nodes | 0.14 |
| Excitation from Chinese words to semantic nodes | 0.14 |
| | |
| Excitation from semantic nodes to English words | 0.14 |
| Excitation from semantic nodes to Chinese words | 0.14 |
| | |
| Input excitation Chinese characters | 0.045 |
| Input excitation English words | 0.11 |
| Input inhibition Chinese characters | -2 |
| Input inhibition English words | -2 |