

ARTICLE

A tale of two frequency effects: Toward a verification model of L2 word recognition

Nan Jiang^{1*}, Man Li¹ and Taomei Guo²

¹University of Maryland and ²Beijing Normal University

*Corresponding author. Email: njiang@umd.edu

(Received 10 January 2019; revised 07 October 2019; accepted 09 October 2019; first published online 25 November 2019)

Abstract

This study examined the activation of first language (L1) translations in second language (L2) word recognition in a lexical decision task. Test materials included English words that differed in the frequency of their Chinese translations or in their surface lexical frequency while other lexical properties were controlled. Chinese speakers of English as a second language of different proficiencies and native speakers of English were tested. Native speakers produced a reliable lexical frequency effect but no translation frequency effect. English as a second language speakers of lower English proficiency showed both a translation frequency effect and a lexical frequency effect, but those of higher English proficiency showed a lexical frequency effect only. The results were discussed in a verification model of L2 word recognition. According to the model, L2 word recognition entails a checking procedure in which activated L2 words are checked against their L1 translations. The two frequency effects are seen to have two different loci. The lexical frequency effect is associated with the initial activation of L2 lemmas, and the translation frequency effect arises in the verification process. Existing evidence for verification in L2 word recognition and new issues this model raises are discussed.

Keywords: lexical decision; second language; translation frequency effect; verification; word recognition

In adult second language (L2) learning, learners tend to associate L2 words with their translations in the first language (L1) in the learning process, particularly in classroom settings and for L2 words that have L1 translations. This raises the question of whether L1 translations play a role in L2 word recognition. Specifically, are L1 translations activated in the process of L2 word recognition? If they are, what role do they play? Is their activation a by-product or an integral part of L2 word recognition?

Bilingual representation and development models recognize the likelihood of L1 translation involvement in L2 word recognition, at least at an early stage of L2 development (e.g., Jiang, 2000; Kroll & Stewart, 1994). In Jiang's three-stage model of L2 lexical development, for example, lexical access in an L2 at the early stage is mediated by the activation of L1 translations, which provides semantic and syntactic information for L2 words. These models also recognize that a transition from

obligatory L1 translation activation to more autonomous L2 word recognition may occur as individuals' L2 proficiency increases and as direct links are established between L2 words and concepts. Such a transition can also be accommodated in other conceptions of L2 processing, such as the lexical entrenchment hypothesis (Brysbaert, Lagrou, & Stevens, 2017; Diependaele, Lemhöfer, & Brysbaert, 2013) and the bilingual interactive-activation developmental model (Grainger, Midgley, & Holcomb, 2010). In the former case, for example, increased experiences in L2 may make L2 lexical representations not only stronger or more precise but more autonomous, such that the processing of L2 words does not have to involve the activation of their L1 translations.

There is ample evidence for the influence or involvement of L1 in L2 use. For example, L2 lexical errors can often be traced back to L1 lexicalization patterns. Thus, an Arabic-speaking learner of English as a second language (ESL) may confuse *bakery* and *oven*, and a Finnish ESL speaker may confuse *tongue* and *language* because these concepts and words are conflated in their L1. Such L1-driven lexical errors are well documented (e.g., Hall & Ecke, 2003; Olshtain & Cohen, 1989; Zughoul, 1991). There is also evidence showing the adoption of L1 processing strategies in L2 word recognition. For example, L2 speakers with a logographic L1 background tended to rely on orthography more and on phonology less than those with an alphabetic L1 in L2 visual word recognition and reading (e.g., Koda, 1989; Wade-Woolley, 1999; Wang & Koda, 2005). Finally, L1 involvement in L2 processing has been demonstrated in research that explored whether lexical access is selective or nonselective in bilingual speakers. For example, in a monolingual English lexical decision task (LDT), the performance of the Dutch–English bilinguals was found to be affected by the number of Dutch lexical neighbors of the English targets (van Heuven, Dijkstra, & Grainger, 1998, Experiment 4). In a study by Jared and Kroll (2001), the naming of L2 English words was affected by some phonological properties of L1 French words (Experiment 4) among French–English bilinguals.

However, L1 involvement is different from L1 translation involvement. L1 influence can occur as a result of cross-linguistic differences in semantic structures (in relation to L2 lexical errors) and processing strategies (in relation to the reliance on phonological vs. orthographic information in L2 word recognition). In both cases, L1 influence can occur without the activation of specific L1 translations. Bilingual selectivity studies are often concerned with the activation of the nontarget language rather than that of specific L1 translations, too. In addition, these studies often reported the activation of L2 in L1 processing (e.g., Colomé 2001; Moon & Jiang, 2012). Thus, these results are less informative about whether L1 translations are activated and participate in L2 word recognition.

Automatic activation of L1 translations in L2 word recognition became a specific topic for investigation in several recent studies. These studies adopted an innovative manipulation of L1 translation overlap first used by Thierry and Wu (2004, 2007). To illustrate this manipulation with examples from Thierry and Wu (2007), pairs of English words were differentiated in terms of two variables: (a) whether they were semantically related or unrelated, and (b) whether their disyllabic Chinese translations shared (or repeated) a Chinese character, thus creating four conditions: related repeated (*post-mail*, *youzheng-youjian*), related unrepeated (*wife-husband*, *qizi-zhangfu*), unrelated repeated (*train-ham*, *huoche-huotui*), and unrelated unrepeated

(*apple-table*, *pingguo-zhuozi*). The critical comparison for the present discussion was unrelated pairs that shared or did not share one of the characters in their Chinese translation (e.g., *train-ham* and *apple-table*). The rationale underlying this manipulation was that if this partial translation overlap affected the performance of Chinese ESL speakers but not English native speakers (NS), one may take this partial translation overlap effect as evidence for automatic translation into their L1 in the process of L2 word recognition.

Thierry and Wu reported a series of three studies in which Chinese ESL speakers and English NS were asked to decide if two English words were related in meaning (2004, 2007; Wu & Thierry, 2010) and both behavioral and electrophysiological data were collected. In the first study (Thierry & Wu, 2004), Chinese ESL speakers responded to the unrelated repeated condition more slowly than the unrelated un-repeated condition, indicating an inhibitory effect of partial translation overlap. Electrophysiological data also showed an elevated N400 for unrelated pairs whose Chinese translations shared a character. Native English speakers showed no such differences. In the second study (Thierry & Wu, 2007), the Chinese ESL participants showed no difference in reaction time (RT) or error rate (ER) between items whose translations did and did not share a Chinese character, but these participants showed a reduced N400 amplitude for test items whose Chinese translations shared a character. Again, English NSs showed no such difference. In the third study, translation repetition was differentiated between sound repetition (the two translations involved two characters that shared the same pronunciation, e.g., *experience-surprise*, *jingyan-jingya*, 经验-惊讶) and spelling repetition (the two translations shared one character that had different pronunciation, e.g., *accountant-conference*, *kuaiji-huiyi*, 会计-会议). Translation repetition did not produce any difference in behavioral data among Chinese ESL speakers, but they showed a reduced N400 for the sound repetition condition. The partial translation overlap effect reported in these three studies provided the first set of evidence for the automatic activation of L1 translations in L2 word recognition with both behavioral and electrophysiological data, even though these data were not always consistent across the studies.¹

Zhang, van Heuven, and Conklin (2011) reported another study that involved a similar manipulation. They tested Chinese–English bilingual speakers on English prime-target pairs such as *east-thing* in a monolingual English LDT. The critical stimuli were pairs of English words whose Chinese translations overlapped. For example, the English word *east* is typically translated into *dong* in Chinese, which is the first character of the disyllabic opaque compound *dongxi*, literally meaning *east west*, but together meaning *thing* as a compound. Thus, to a Chinese–English bilingual speaker, *east* and *thing* are linked through a chain of connections from *east* to *dong* to *dongxi* to *thing*, even though these two words are not semantically or associatively related. Under a masked priming condition with an SOA of 80 ms (57 ms for prime duration and 23 ms for backward mask), they were able to observe reliable priming effects for prime-target pairs such as *east-thing* or *thing-east*. This priming effect was interpreted to have arisen from the automatic activation and involvement of L1 translations, conceivably through a chain activation sequence such as *east-dong-dongxi-thing*. However, this finding was not replicated in a subsequent study involving the same materials and the task (Wen & van Heuven, 2018).

Similar findings were reported in three studies involving bimodal bilinguals. Participants in these studies were deaf bimodal bilinguals who used either American Sign Language or Spanish Sign Language (*lengua de signos española*) as their native languages and English or Spanish as their L2. They were asked to perform a semantic judgment task on pairs of words in their L2 English or Spanish. Similar to the manipulation adopted by Thierry and Wu (2007), the stimuli were selected that had or did not have form overlap in their sign language translations. In both Morford, Wilkinson, Villwock, Piñar, and Kroll (2011) and Villameriel, Dias, Costello, and Carreiras (2016), semantically related word pairs with American Sign Language form overlap were responded to faster than pairs without overlap, and for semantically unrelated pairs, items with overlap were responded to more slowly than pairs without overlap. In Meade, Midgley, Schyr, Holcomb, and Emmorey (2017), L2 pairs with translation overlaps produced a smaller N400 than those without overlaps. Unrelated pairs with translation overlap also produced slower RTs than those without overlap. All these findings provided further evidence for automatic L1 translation activation in L2 word recognition. It should be noted that Villameriel *et al.* also included a group of hearing bimodal bilinguals who learned Spanish Sign Language as an L2 and found similar sign language activation in processing their L1 Spanish.²

The studies reviewed above represented recent efforts to directly assess the activation of L1 translations in L2 word recognition. They also shared the basic approach of manipulating translation overlap. While being an innovative manipulation, this approach has potential limitations. The first one is that the partial translation overlap effect assessed in this manipulation may not be robust enough to provide consistent results, as it requires the spreading of activation across multiple links, first between L2 words and its L1 translations, and between two L1 translations, and then between L1 translations and L2 words. Assuming that activation decays over successive links, activations involving multiple links may not always be reliably detected. This may explain the inconsistencies in the results in Thierry and Wu's three studies, in the behavior results of Thierry and Wu (2001) and Meade *et al.* (2017), and in the results between Zhang *et al.* (2011) and Wen and van Heuven (2018).

Second, even if the translation overlap effect may be assessed and interpreted as evidence for the automatic activation of L1 translations in L2 word recognition, it says very little about what role L1 translations play in the process. The activation of L1 translations may be an integral part of L2 word recognition, or a by-product of L2 word recognition, the latter of which seemed to be endorsed by Thierry and Wu (2007). They suggested that L1 translation activation found in their study was post-lexical. Their conclusion was based on a difference they found between Chinese–English bilinguals and monolingual Chinese speakers. The latter group was tested with the Chinese translations of the English stimuli in the same task, and they showed a P2 effect. It was a greater positive-going amplitude for repeated than for nonrepeated pairs at around 200 ms after the onset of the stimuli and is often interpreted as reflecting repetition priming. In contrast, Chinese–English bilinguals did not show a P2 while processing English stimuli.

Finally and more importantly, even when a reliable translation overlap effect is found, it may not be unequivocally interpreted as evidence for the involvement of L1

translations. An alternative interpretation was offered by Costa, Pannunzi, Deen, and Pickering (2017). They suggested that an association may be established between two otherwise unrelated L2 words because their L1 translations are linked in some way. For example, *train* and *ham* may become linked directly in the mind of a Chinese ESL speaker because their Chinese translations, *huoche* and *huotui*, share the first character. The presence of such lexical links between L2 words can produce a partial translation overlap effect as reported by Thierry and Wu (2004, 2007) without resorting to L1 translation activation (but see Oppenheim, Wu, and Thierry, 2018, for a response and related findings).

The purpose of the present study was to examine L1 translation involvement in L2 word recognition with a different manipulation. We asked Chinese–English bilingual speakers to perform a LDT on English words that differed in the frequency of their Chinese translations. Two words of similar frequency in one language may have translations in another language that differ significantly in frequency due to linguistic and cultural reasons.³ For example, both *revolution* and *signature* have a frequency of 12 occurrences per million in Brysbaert and New (2009). However, their Chinese translation, *geming* and *qianming*, had a frequency of 1,963 and 3 per million, respectively, according to a Chinese Frequency Dictionary published by Beijing Language and Culture University (BLI, 1986).

This manipulation of translation frequency offers an opportunity to examine the role of L1 translation in L2 word recognition. If L1 translations are not involved in L2 word recognition, or if their activation is a by-product of L2 word recognition, the frequency of L1 translations should not affect L2 word recognition. Thus, when the test materials are properly controlled, both English NS and Chinese ESL speakers should show no difference for words with higher and lower frequency translations. However, if the activation of L1 translations is an integral part of L2 word recognition, we would expect to observe an L1 translation frequency effect in that Chinese–English bilinguals, but not English NS, would respond to English words with higher frequency Chinese translations faster than words with lower frequency translations.

This manipulation has at least three advantages. First, the observation of an L1 translation frequency effect does not require the activation of multiple lexical links. In this sense, it is a more direct or robust manipulation for assessing the involvement of L1 translations. Second, because an L1 translation frequency effect was assessed without the involvement of two L2 words whose L1 translations overlapped, Costa et al.'s (2017) alternative explanation would be no longer relevant. Thus, the effect, if obtained, offers more unequivocal evidence for L1 activation in L2 word recognition. Third and finally, a translation frequency effect does not only indicate the activation of L1 translations in L2 word recognition. The finding that the lexical properties of L1 translations affected L2 word recognition would mean that the activation of L1 translations is an integral part of L2 word recognition, rather than a by-product of the latter.

Two experiments are reported here. Experiment 1 represented our initial attempt to examine the translation frequency effect. Experiment 2 was intended to replicate the results of Experiment 1 and explore two additional issues: the relationship between the translation frequency effect and the lexical frequency effect and the effect of long immersion on L2 lexical processing.

Experiment 1

Method

Participants

Two groups of Chinese ESL speakers and a group of English NS participated in the study. The first group of ESL participants were 34 college students recruited in China. They represented ESL speakers with minimal immersion experience in the target language. A second group of ESL speakers were recruited in the United States. These 28 participants were graduate students and visiting scholars studying at an American university. Their length of residence in the United States varied from 1 month to 45 months, with an average length of residence of 15 months. Differentiating participants in L2 proficiency in an experimental setting has long been a thorny issue (Grosjean, 1998). We used the length of immersion as an indication of L2 proficiency in both experiments, as length of stay in the target language has been found to differentiate L2 speakers in lexical processing (e.g., Athanasopoulos, 2007; Athanasopoulos, Dering, Wiggett, Kuipers, & Thierry, 2010; Cook, Bassetti, Kasai, Sasaki, & Takahashi, 2006). A group of 34 native English speakers studying at the same university as the second ESL group served as controls. NS speaking participants received course credit, and the ESL group (non-native speakers; NNS) received \$10 for their participation.

Materials

The critical stimuli for this study consisted of two sets of English words that were matched for frequency and length but differed in the frequency of their Chinese translations. Developing such materials can be a challenge as the frequency of L2 words and that of their L1 translations are likely to be correlated (e.g., Wen & van Heuven, 2017). Several steps were taken in the process. We started with 140 high-frequency Chinese nouns and verbs based on BLI (1986). These words were translated by the first author into English and the frequency of the English translations were identified using Brysbaert and New (2009). We then selected 80 pairs of translations that had the greatest difference in frequency between the two members. For example, *tongzhi* (comrade) had a frequency of 1,825 per million, but its English translation has a surface frequency of 9 per million. The 80 English words were then given to three Chinese ESL speakers who were asked to translate them into Chinese to make sure the selected English words had a unique Chinese translation. Sixty-one English words were translated into the same Chinese word by all three informants and were retained for the next step. For each of the 61 words, 2 additional English words were identified that had the same part of speech and similar frequency and length. These 122 English words were then translated into Chinese by the first author, and the frequency of these Chinese words was identified. A set of 40 English words that had lower frequency Chinese translations and were likely to have a unique Chinese translation were selected and given to three additional Chinese ESL speakers. They were asked to translate these English words into Chinese to confirm their unique Chinese translation.

Based on these procedures, 32 English words were selected that met the following criteria: (a) they each had a unique disyllabic Chinese translation; (b) they were all nouns or verbs; (c) 16 of the English words had higher frequency Chinese

Table 1. Examples and lexical properties of the test materials for the higher translation frequency (HTF) and lower translation frequency (LTF) conditions in Experiment 1

Examples	Mean frequency (per million)	Mean length (number of letters)	Mean familiarity rating scores	Chinese translation	Mean Chinese translation frequencies (per million)
HTF <i>research</i>	21.6	7.5	7.8	<i>yanjiu</i> 研究	810.3
LTF <i>memory</i>	22.0	7.5	7.5	<i>jiyi</i> 记忆	26.9

translations (average frequency being 810.3 per million), and the other 16 had lower frequency Chinese translations (average frequency being 26.9 per million); and (d) these two sets of English words were matched for frequency (21.6 and 22.0 per million) and length (both sets 7.5 letters). To ensure that the corpus-based frequency information for the English words reflected the degree of familiarity of these words to Chinese ESL speakers and to ensure that these words were familiar to Chinese ESL speakers in China, a group of 16 participants in China were asked to rate their familiarity with these English words after they were tested. They did so on a 1–10 scale with 1 indicating minimum familiarity and 10 indicate highest degree of familiarity. The two sets of English words received an average rating scores of 7.8 and 7.5, respectively, and the difference was not significant ($t = 1.27, p = .24$). The lexical properties and examples of the stimuli are shown in Table 1.

The two sets of English words and their Chinese translations can be found in Appendix A. The two sets of English words, along with 18 filler items, which were English nouns and verbs of varying frequencies, and 50 nonwords were used to construct a single presentation list of 100 items.

Before the main experiment with English words was conducted, a pretest was done to check if the Chinese translations themselves could produce a reliable frequency effect. Six Chinese NSs were asked to perform a LDT on the higher and lower frequency Chinese words. They produced a significant frequency effect of 45 ms (507 ms vs. 552 ms; $t = 2.8, p < .05$), confirming the effectiveness of the frequency manipulation.

Procedure

Participants were tested individually. They were asked to decide whether a letter string was an English word or not and to respond as quickly and accurately as possible. Each test item began with a fixation point presented for 500 ms, which was then followed by a letter string. The stimuli remained on the screen until a response was provided. Test items were randomized for each participant. Ten practice items preceded test items. The test lasted for less than 20 minutes for most participants. Stimulus presentation and data collection were done with DMDX (Forster & Forster, 2003) on a PC.

Results and discussion

RT and ER data were collected from each participant. Any RT that was associated with an incorrect response was excluded from analysis, so was RT that lied outside

Table 2. The mean RT and ER for higher and lower translation frequency (HTF, LTF) L2 words from the three groups of participants

Conditions	NS <i>n</i> = 34		NNS1 <i>n</i> = 34		NNS2 <i>n</i> = 28	
	RT	ER	RT	ER	RT	ER
HTF <i>n</i> = 16	580 (109)	1.5 (2.7)	706 (120)	3.3 (5.2)	704 (154)	3.4 (4.7)
LTF <i>n</i> = 16	655 (82)	3.4 (6.4)	756 (141)	11.4 (11.0)	762 (194)	3.2 (4.9)
<i>Difference</i>	-15	1.9	50*	8.1**	58*	-0.2

Note: *Significant at $p < .05$ in participant analysis. **Significant at $p < .05$ in both participant and item analyses.

of the cutoff range set at 300 ms and 3000 ms, or two standard deviation of the mean of the same participant, which affected 4.1%, 4.5%, and 5.3% of the data for the three groups. The remaining data were used to calculate the participant and item means for RT and ER for each condition, which were used in by-participant and by-item statistical analysis. Raw RT data were used in mixed-effects regression analysis after trimming and transformation to improve normal distribution. Table 2 displays the participant RT and ER means for the two conditions.

Analyses carried out using the SPSS general linear model-repeated measures procedure included the participant group as the between-participant factor and translation frequency as a within-participant factor. There was a main effect of participant group in both RT and ER in both the participant and item analyses, RT: $F_1(2, 93) = 16.0, p < .05, \eta^2 = .25, F_2(2, 45) = 70.2, p < .05, \eta^2 = .75$; ER: $F_1(2, 93) = 10.4, p < .05, \eta^2 = .18, F_2(2, 45) = 6.2, p < .05, \eta^2 = .21$. There was also a main effect of translation frequency in RT and ER in both analyses, RT: $F_1(2, 93) = 15.1, p < .05, \eta^2 = .24, F_2(2, 45) = 4.1, p = .049, \eta^2 = .08$; ER: $F_1(2, 93) = 13.0, p < .05, \eta^2 = .12, F_2(2, 45) = 6.29, p < .05, \eta^2 = .12$. More importantly, there was also a significant interaction between participant group and translation frequency in RT in participant analysis and in ER in both analyses, RT: $F_1(2, 93) = 15.17, p < .05, \eta^2 = .21, F_2(2, 45) = 2.15, p = .12, \eta^2 = .08$; ER: $F_1(2, 93) = 7.50, p < .05, \eta^2 = .13, F_2(2, 45) = 3.88, p < .05, \eta^2 = .14$.

Separate analyses of the three groups of participants' performance showed that there was a marginally significant difference between higher and lower translation frequency conditions in RT in participant analysis among NSs, RT: $t_1 = 1.8, p = .075, t_2 = 0.97, p > .05$, but the difference was in the opposite direction from that of the translation frequency effect. There was no difference in their ER data, ER: $t_1 = 1.5, p > .05, t_2 = 1.1, p > .05$. However, both groups of ESL speakers showed a significant difference in RT favoring the higher frequency condition, at least in the participant analysis, NNS1: $t_1 = 4.7, p < .05, t_2 = 2.0, p = .06$; NNS2: $t_1 = 4.7, p < .05, t_2 = 1.4, p = .15$. The first NNS group also showed a significant difference in ER, NNS1: $t_1 = 4.0, p < .05, t_2 = 2.6, p < .05$; NNS2: both $ts < 1$.

Analysis of the two ESL groups' data showed that they were similar in RT, both $Fs < 1$, but the two groups were significantly different in ER, as can be seen in a main effect of participant group, $F_1(1, 60) = 8.9, p < .05, \eta^2 = .12, F_2(1, 30) = 5.4, p < .05, \eta^2 = .15$, as well as a significant interaction of translation frequency and participant group in ER in both analyses, $F_1(1, 60) = 11.1, p < .05, \eta^2 = .15, F_2(1, 30) = 5.8, p < .05, \eta^2 = .16$.

Mixed-effects linear regressions of RT data were also run that provided confirmation for the analysis of variance (ANOVA) results. In these analyses, the same data trimming method was used as in the ANOVA. Reciprocal transformation of the raw data was done to improve normal distribution. *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2016, 2017; also see Luke, 2017), a package developed for the R environment, was used for obtaining *p* values in the analyses where item and participants were treated as random variables and translation frequency as a fixed variable for each group. Among the three groups, NS showed no significant difference between the two conditions, but both NNS groups showed marginally significant effects, NS: $t = -0.76$, $p = .45$; NNS1: $t = 1.84$, $p = .07$; NNS2: $t = 1.67$, $p = .10$.

Among the results from Experiment 1, the most important finding was a translation frequency effect for both NNS groups while the NS participants showed an opposite pattern. NNS participants were faster in responding to English words with higher frequency Chinese translations than to words with lower frequency translations. This finding suggested that L1 translations were activated and their activation was more than a by-product of L2 word recognition. Instead, they affected L2 word recognition.

Before we discuss the specific nature of this effect, we hoped to replicate this finding and explored two additional issues. First, we wanted to know if a lexical frequency effect can also be observed independently of their translation frequency and in the same experimental context where a translation frequency effect is obtained. Frequency effects in L2 word recognition have been well documented in previous research (e.g., de Groot, Borgwaldt, Bos, & Van den Eijnden, 2002; Jiang, 1999; Ko, Wang, & Kim, 2011), but as higher frequency L2 words may tend to have higher frequency L1 translations (e.g., Wen & van Heuven, 2017), it is not clear whether the observed lexical frequency effect in previous studies reflected the influence of lexical frequency or translation frequency. Lexical frequency was separated from translation frequency in Experiment 2 by developing test materials that were different in lexical frequency but matched for translation frequency. Second, we hoped to know whether this L1 translation involvement will continue among L2 speakers who have reached a steady state in an L2. This was done by testing Chinese ESL participants who had lived in North America for an extended period of time when they had to use English for their work.

Experiment 2

Method

Participants

The participants in Experiment 2 included one group of 37 English NS and two groups of Chinese ESL speakers. The first NNS group, referred to as the NNS1 group, included 35 Chinese-speaking students who were studying at the University of Maryland at the time of testing. Most of them were graduate students, but a small number were undergraduates. A majority of them had a TOEFL score of more than 100, but 4 had a score between 90 and 100. They were almost all in their 20s, with the exception of a few in their early 30s. These participants were from the

same population as those in the NNS2 group in Experiment 1. There were 29 Chinese ESL speakers in the second NNS group (NNS2). To qualify as a participant in this group, an individual had to have obtained a graduate degree in a Canadian or US university, used English as their work language, and lived in North America for a minimum of 10 years. All but 1 participant in this group held a PhD from a US or Canadian university. The remaining 1 had a MA from a US institution. All but 4 participants were university professors, with the other 4 being corporate or government employees. A majority of them had lived in the United States for 15 years or more at the time of testing. They were also much older than the NNS1 group, with a majority of them approaching or in their 40s. This long-immersion group represented L2 speakers who have reached a steady state in their L2 English proficiency at least as far as lexical processing was concerned.

Materials and procedure

Test materials were developed to manipulate translation frequency and lexical frequency separately. The same 32 English words with higher and lower frequency Chinese translations from Experiment 1 were used to assess the translation frequency effect. Another 32 English words were selected that differed in lexical frequency (based on Brysbaert & New, 2009) but matched for length and translation frequency. Examples of the test materials in these four conditions along with their lexical properties are shown in Table 3. The actual test materials can be found in Appendix A. All 64 word trials and 64 nonword trials formed a single list, which was randomized individually for each participant. The procedure was the same as in Experiment 1.

Results

Data trimming and statistical analyses followed the same procedures as in Experiment 1. The former affected 4.3%, 4.4%, and 5.6% of the data for the NS, NNS1, and NNS2 groups. The mean RT and ER for the four conditions from the three groups are listed in Table 4.

In analyzing data, the translation frequency effect and lexical frequency effect were analyzed separately. In each of these two sets of analyses, ANOVA was first performed on RT and ER means in which frequency (translation or lexical) was treated as a within-participant variable with two levels (high, low), and the participant group was treated as a between-participant variable with three levels (NS, NNS1, NNS2). Planned comparisons were then performed for each participant group using paired-samples *t* tests in SPSS for all means. Linear regressions were then run on RT data for each participant group using the *lmerTest* package in the R environment.

The lexical frequency effect

Analysis of variance of the data from the lexical frequency manipulation produced a main effect of participant group in RT, $F_1(2, 98) = 17.0$, $p < .05$, $\eta^2 = .26$, $F_2(2, 45) = 78.1$, $p < .05$, $\eta^2 = .77$, but not in ER, $F_1(2, 98) = 2.3$, $p > .05$, $F_2(2, 45) = 1.85$, $p > .05$. The NS group was the fastest, and the NNS2 group was

Table 3. Examples and lexical properties of the test materials for the higher translation frequency (HTF), lower translation frequency (LTF), higher lexical frequency (HLF), and lower lexical frequency (LLF) conditions in Experiment 2

Conditions	Examples	Mean frequencies (per million)	Mean lengths (number of letters)	Chinese translations	Mean translation frequencies (per million)
HTF	<i>research</i>	22.1	7.5	研究	810.3
LTF	<i>evidence</i>	23.9	7.5	证据	27.8
HLF	<i>room</i>	307.8	6.3	房间	76.3
LLF	<i>carpet</i>	21.6	6.6	地毯	77.8

Table 4. The participants' mean reaction times (RT in millisecond, ms) and error rates (ER, in percentage) for the higher lexical frequency (HLF), the lower lexical frequency (LLF), the higher translation frequency (HTF), and the lower translation frequency (LTF) English target words (standard deviations in parentheses) in Experiment 2

Conditions	NS <i>n</i> = 37		NNS1 <i>n</i> = 35		NNS2 <i>n</i> = 29	
	RT	ER	RT	ER	RT	ER
HLF <i>n</i> = 16	538 (84.1)	2.0 (4.2)	593 (89.2)	1.5 (3.2)	666 (118.2)	1.2 (2.6)
LLF <i>n</i> = 16	558 (80.3)	2.8 (4.2)	622 (85.4)	6.8 (7.9)	715 (147.9)	3.5 (4.9)
<i>Difference</i>	20*	0.8	29*	5.3*	49*	2.3**
HTF <i>n</i> = 16	557 (90.0)	3.4 (5.7)	624 (97.7)	4.5 (4.7)	742 (164.9)	3.4 (5.1)
LTF <i>n</i> = 16	561 (88.1)	4.9 (6.5)	650 (99.0)	5.1 (7.5)	764 (171.5)	3.6 (5.1)
<i>Difference</i>	4	-1.2	26*	0.6	22	-0.2

Note: *Significant at $p < .05$ in paired-samples *t* tests in participant analyses. **Significant at $p < .05$ in paired-samples *t* tests in both participant and item analyses.

the slowest.⁴ There was also a main effect of lexical frequency in RT, $F_1(1, 98) = 3.92$, $p < .05$, $\eta^2 = .07$, $F_2(1, 45) = 11.5$, $p < .05$, $\eta^2 = .20$, as well as in ER, $F_1(1, 98) = 21.8$, $p < .05$, $\eta^2 = .18$, $F_2(2, 45) = 9.43$, $p < .05$, $\eta^2 = .17$. The interaction between the two variables in RT was reliable in participant analysis, $F_1(2, 98) = 3.92$, $p = .050$, $\eta^2 = .07$, $F_2(2, 45) < 1$, so was the ER data, $F_1(2, 98) = 5.50$, $p < .05$, $\eta^2 = .10$, $F_2(2, 45) = 2.30$, $p > .05$.

Planned comparisons were performed with paired-samples *t* tests for the three groups of participants comparing their RT and ER for the high- and low-lexical frequency conditions. Mixed-effects linear regressions of RT data were also done using the R package lmerTest. These results are shown in Table 5.

As is clear from the above analyses and from Table 5, all three groups showed a lexical frequency effect in RT. They responded to high lexical frequency words faster than to low lexical frequency words. Note that this lexical frequency effect was obtained with words that were matched for translation frequency, so it was independent of the latter. In addition, the two NNS groups also showed a lexical frequency effect in ER. They made more errors while responding to lower frequency words than to higher frequency words.

Table 5. Results of planned comparisons of the three participant group's RT and ER and mixed-effects models for the RT data for the lexical and translation frequency manipulation

Independent variable	Group	Data type	Paired-samples <i>t</i> tests						Mixed-effects linear regression	
			Participant analysis			Item analysis			<i>t</i>	<i>p</i>
			<i>t</i> ₁	<i>df</i>	<i>p</i>	<i>t</i> ₂	<i>df</i>	<i>p</i>		
Lexical frequency	NS	RT	-3.16	36	.00	-1.98	15	.06	2.33	.02
		ER	-0.90	36	.37	-0.78	15	.44		
	NNS1	RT	-4.82	34	.00	-1.95	15	.07	2.11	.04
		ER	-4.28	34	.00	-2.47	15	.02		
	NNS2	RT	-4.70	28	.00	-2.22	15	.04	2.14	.03
		ER	-2.45	28	.02	-1.64	15	.12		
Translation frequency	NS	RT	-0.73	36	.46	-0.66	15	.51	1.10	.27
		ER	-1.57	36	.12	-0.96	15	.34		
	NNS1	RT	-3.91	34	.00	-1.77	15	.09	2.13	.04
		ER	-0.45	34	.65	-0.14	15	.88		
	NNS2	RT	-2.48	28	.01	-1.10	15	.28	1.17	.25
		ER	-0.19	28	.84	-0.25	15	.80		

The translation frequency effect

Analysis of variance of the data obtained from the 32 English words that differed in translation frequency produced a main effect of participant group in RT, $F_1(2, 98) = 21.9$, $p < .05$, $\eta^2 = .31$, $F_2(2, 45) = 137.8$, $p < .05$, $\eta^2 = .86$, but not in ER, both $F_s < 1$. There was also a main effect of translation frequency in RT, $F_1(1, 98) = 18.8$, $p < .05$, $\eta^2 = .16$, $F_2(1, 45) = 3.61$, $p = .06$, $\eta^2 = .07$, but not in ER, $F_1(1, 98) = 1.28$, $p > .05$, $F_2(2, 45) < 1$. Significantly, the interaction between the two variables in RT was reliable in participant analysis, $F_1(2, 98) = 3.09$, $p = .050$, $\eta^2 = .06$, $F_2(2, 45) < 1$. There was no significant interaction in ER, both $F_s < 1$.

Planned comparisons were performed with paired-samples *t* tests for the three groups of participants comparing their RT and ER for the high- and low-translation conditions, which was supplemented by mixed-effects linear regression analyses. The results of these analyses are shown in Table 5.

In contrast to the results for lexical frequency, the three groups of participants showed a different pattern in responding to words of higher and lower translation frequency. NSs showed no translation frequency effects, which indicated that the materials were adequately matched for lexical properties other than translation frequency. Among the two NNS groups, the NNS1 group showed a translation frequency effect in RT in both paired-samples *t* tests and in regression analysis. However, the NNS2 group produced a significant difference in participant analysis only in paired-samples *t* tests, with no reliable difference in regression analysis. None of the participant groups produced any significant difference in ER data.

General discussion

Two frequency effects were observed independently of each other in the study. When pairs of L2 words were matched for lexical frequency, L2 speakers, but not L1 speakers, responded to English words with higher frequency L1 translations faster than those with lower frequency translations in a LDT. We refer to this finding as the translation frequency effect. At the same time, when pairs of L2 words were matched for translation frequency but varied in lexical frequency, both NSs and NNSs produced a lexical frequency effect in that they responded to words of higher frequency faster than those of lower frequency.

Lexical frequency effects have been well documented in L2 (e.g., de Groot et al., 2002; Jiang, 1999; Ko et al., 2011). However, as L2 lexical frequency is often correlated with L1 translation frequency (see Wen & van Heuven, 2017), the lexical frequency effect can be a true lexical frequency effect, a translation frequency effect, or a combination of both. By controlling translation frequency, this study was the first to demonstrate that the L2 lexical frequency effect is independent of translation frequency.

This study was also the first to report a translation frequency effect. This finding is consistent with those of Thierry and Wu (2004, 2007; Wu & Thierry, 2010), Morford et al. (2011), Villameriel et al. (2016), and Meade et al. (2017), in suggesting that L1 translations are automatically activated in L2 word recognition. Furthermore, the finding that translation frequency affected L2 word recognition performance suggests that the activation of L1 translations is more likely to be an integral part of L2 word recognition rather than its by-product. At the same time, the lack of this translation frequency effect among the long-immersion participants suggests that L2 words may be accessed autonomously as an individual's L2 experience and proficiency increase.

The findings of the two frequency effects raised two questions. First, what exactly is the role of L1 translations in the process of L2 word recognition? Second, how do these two frequency effects relate to each other? We discuss these two issues in the framework of a verification model of L2 word recognition.

Toward a verification model of L2 word recognition

Verification is a long-standing concept in the study of word recognition. The concept gained its prominence first in the development of models of visual word recognition. In several such models, such as Becker's (1976) verification model, Paap, Newsome, McDonald, and Schvaneveldt's (1982) activation-verification model, and Forster's (1976) search model, word recognition is seen as proceeding in two steps: that of visual input analysis and verification. The first step (e.g., the sensory-feature-extraction process in Becker's model) generates a set of candidates based on the analysis of visual input. The second step (referred to as the verification process in Becker's model or the post-access check in Forster's model) checks these candidates with the visual input to ensure a good match and to isolate a single candidate. Similarly, in Levelt, Roelofs, and Meyer's (1999) model of lexical access, an activated lemma will be checked against the related concept in a verification procedure before it is finally selected.

The concept of verification is also prevalent in the discussion of L2 or bilingual processing, even though what is verified differs in different conceptions. For example, in Grainger's (1993) bilingual activation verification model, which was intended to explain nonselective lexical access in bilinguals, visual stimuli may initially activate all words in both languages that match the stimuli at least partially. A verification process then follows in which language context information is used to determine which activated words should be checked against the input first. For example, in an LDT involving a mixture of French and English words, if the preceding item is an English word, then the English candidates will be checked first. In Green's (1998) inhibitory control model of bilingual production, lexical production in a task such as picture naming entails a checking procedure to make sure that the activated lemma is linked to the concept with the right language tag. This checking procedure allows a bilingual speaker to name the picture in the right language, even though a picture may activate lemmas in both languages.

In our conception of L2 word recognition, initial analysis of the input helps identify an L2 word that best matches the input. This is a process shared between L1 and L2 word recognition. However, L2 word recognition entails a verification or checking procedure whereby an activated L2 word is checked against its L1 translation. If an L1 translation is found for the word, the identity of the L2 word is confirmed. In an LDT, this means a decision is made only after an L1 translation is activated. Within this conception, the lexical frequency effect and the translation effect have different loci. The former results from the first step of L2 word recognition. Words of higher frequency are accessed faster than those of lower frequency because activation accumulates faster, the resting level is higher, or the candidates become available for subsequent checking sooner for high-frequency words, depending on what model of word recognition is adopted. The translation frequency effect, in contrast, arises from the verification process. Higher frequency L1 translations may be located or activated more easily or faster for the checking procedure than lower frequency translations, thus leading to the translation frequency effect.

Several findings corroborate with the translation frequency effect in suggesting the presence of a verification process in L2 word recognition. The first one is the translation lexicality effect in L2 compound processing reported by Wang and colleagues (Cheng, Wang, & Perfetti, 2011; Ko, Wang, & Kim, 2011; Wang, Lin, & Gao, 2010). They tested Chinese and Korean ESL speakers on English compound words such as *classroom* and *honeymoon* in a LDT. In addition to the semantic transparency manipulation as illustrated by these two examples, they also included a translation lexicality variable by using English compounds that were and were not directly translatable into Chinese or Korean. For example, the English compound *toothbrush* was classified as directly translatable because its Chinese translation *yashua* (*ya* = *tooth*, *shua* = *brush*) shared the same two component morphemes. In contrast, a direct morpheme-based translation of the English compound *schoolbook* would produce an illegitimate Chinese word *xiaoshu*. Its correct translation is *keben*, which literally means *lesson book*, so *schoolbook* is not directly translatable. In all three studies, L2 speakers were found to respond to translatable L2 compounds faster or more accurately than to untranslatable ones. While these authors interpreted their results in terms of morphological decomposition, the verification model offers an alternative and more specific explanation. In performing an LDT on L2

compounds, an L2 speaker routinely activates its L1 translation in a verification process. When an L2 compound and its L1 translation have complete morphemic overlap, for example, *toothbrush-yashua*, an L2-L1 lexical link is easier to build and consolidate in the learning process. Their stronger L2-L1 links means L1 translations are more readily available for verification, thus facilitating L2 speakers' performance on these words. In contrast, untranslatable compounds such as *textbook* are likely to have weaker L2-L1 links. Thus, L1 translations may take longer to activate in the verification process. This explanation is more direct in the sense that decomposition itself is not sufficient to explain the finding. Verification with L1 translation is the real reason for the observation of the effect.

Another translatability effect was reported by Gollan, Montoya, Fennema-Notestine, and Morris (2005), who asked bilinguals to name pictures. Among the variables considered in the study was translatability, which was determined by considering how many participants were able to provide a correct translation for a target, which produced 121 highly translatable picture names and 58 less translatable ones. The results showed that both monolingual and bilingual participants showed a translatability effect, responding to pictures with translatable names faster than less translatable names, but more important, bilinguals showed a significantly stronger translatability effect than monolinguals. The authors explained the monolinguals' performance by attributing it to the fact that translatability was confounded with other lexical properties such as frequency, familiarity, and length. They used the increased translatability benefit among bilinguals as evidence to argue against cross-language interference, but did not provide an explanation of this bilingual advantage. Within the conception of the verification model, this increased translatability benefit is predictable for bilinguals. In performing a picture naming task, a bilingual speaker may perform a verification procedure when a picture name is activated. For items whose names were more translatable, their translations are more readily available for the verification process, thus facilitating the naming task. The verification process may take longer for items whose translations were less available.

Finally, the translation ambiguity effect in L2 word recognition provided further evidence for verification. Degani, Prior, Eddington, Arêas da Luz Fontes, and Tokowicz (2013, discussed in Tokowicz, 2014) tested Spanish-English bilinguals in an English LDT. The stimuli included words with a single unique or multiple Spanish translations (the translation ambiguity manipulation). Even though it was a monolingual English LDT, the bilingual participants were found to respond to English words with multiple translations less accurately. The finding can be explained rather straightforwardly within the verification model. The verification process for L2 words with a single translation is simple and straightforward. However, when multiple L1 translations are available, confusion or competition may arise in verification, thus producing the observed translation ambiguity effect.

An inevitable question in this discussion is why such a verification process should occur in L2 word recognition. An answer may lie in L2 learners' tendency in associating L2 words with their L1 translation in L2 learning. Note that an L2 is often learned after the establishment of an L1 lexicosemantic system. In contrast to the simultaneous development of lexical form and meaning (or words and concepts) in L1 learning, L2 words are often learned without substantial concurrent semantic development. Instead, L2 learners rely on their L1 lexicosemantic system for

form–meaning mapping in vocabulary acquisition. As a result, a close association between L2 words and their L1 translations is formed and relied on at an early stage of L2 learning. This results in the activation of L1 translation in L2 word recognition, a phenomenon well recognized in Kroll and Stewart's (1994) model of bilingual representation, in Jiang's (2000) model of L2 lexical development, and in Ecker's (2015) parasitic model of vocabulary acquisition.

A few issues associated with verification

The proposal of a verification process also raises several questions that are yet to be explored. The first one is related to L2 proficiency. Will this verification process stay with highly proficient L2 speakers or can it be bypassed at some point? The length-of-immersion manipulation in Experiment 2 was intended to explore this issue. The two NNS groups of different immersion length did show a different pattern on translation frequency manipulation. The shorter immersion group (NNS1) showed a robust translation frequency effect in RT in both *t* tests and mixed-effects models, but the longer immersion group (NNS2) showed a reduced effect, which was no longer reliable in mixed-effects linear regression analysis. This group difference may indicate a transition where L1 translations become less involved in L2 word recognition.

A similar proficiency-based difference in L1 translation activation in L2 processing was also reported by Fitzpatrick and Izura (2011). They tested Spanish–English bilinguals in two tasks, a word association (WA) task in both L1 Spanish and L2 English followed by an LDT in Spanish. The test materials in the LDT included the L1 Spanish words whose English translations were used as stimulus words in the WA task and Spanish words whose translations were not used in the WA task. They found that the participants responded to L1 Spanish words faster by 27 ms in the LDT if their English translations were used in the WA task. They suggested that while these participants performed the WA task in L2, the L1 translations were automatically activated, resulting in facilitation in the subsequent L1 LDT. More importantly, this effect was reliable only for the less proficient group. Similar to the findings of the present study, the high-proficiency group showed a difference in the same direction, but the difference of 15 ms was not statistically reliable. These results suggest that this verification may be eventually bypassed as one's L2 proficiency increases, but more research is needed to confirm these findings and explore the details regarding the relation between verification and L2 proficiency.

The second question is related to word type, specifically to L2 words without L1 translations. Several studies reviewed earlier already demonstrated that L2 word recognition or production was affected by the extent to which an L2 can be translated into an L1 word (Cheng *et al.*, 2011; Degani *et al.*, 2013; Gollan *et al.*, 2005; Ko *et al.*, 2011). However, little is known about what happens to L2 words that do not have an L1 translation. One may speculate that these words will take longer to process, as they have to rely on a time-out mechanism to terminate the verification process when an L1 translation cannot be found after some time. This speculation is based on the assumption that verification is a universal process applicable to all L2 words. However, it may also be suggested that such words would be processed faster, on the

assumption that the language processor has learned not to adopt a verification procedure while processing these words. If an L2 word does not have an L1 translation, it means the concept underlying the lexical form is not present or lexicalized in the learner's L1. As a result, concurrent semantic development has to accompany word learning, just like learning a word in the L1. For this reason, these words may take longer to learn. However, once learned, the lexical form is mapped onto semantic content that is specific and closely linked to this L2 word. In lexical access, this means the translation verification process can be bypassed, resulting in a faster processing time. This is certainly an interesting issue yet to be explored.

Additional questions concern language processing tasks and bilingual speakers. A lexical or semantic task, for example, LDT or semantic relatedness judgment, was adopted in the studies that demonstrated L1 translation involvement (e.g., Cheng et al., 2011; Ko et al., 2011; Thierry & Wu, 2004, 2007). These tasks require deeper lexical processing in comparison to a task such as word naming, which may be affected by lexical properties (e.g., phonological regularity) different from those in other tasks, and thus sometimes producing different lexical processing results (e.g., Andrews, 1982; Colombo, Pasini, & Balota, 2006; Kim & Davis, 2003; Perea & Carreiras, 1998). Whether verification is as important in naming as in LDT, or more generally, whether verification is universal or task specific, is yet to be explored. In addition, if we are right in suggesting that the obligatory verification process is a result of learning an L2 by associating L2 words with L1 translations, verification should be less likely to occur for simultaneous bilinguals or in L1 word recognition among bilinguals. If L2 translations are also involved in L1 word processing, one needs to explore whether such translation activation is similar in nature or plays different roles.

Conclusion

We interpret the finding of two independent frequency effects as evidence for the presence of a verification procedure in L2 word recognition whereby L1 translations are checked against the activated L2 words. The verification model outlined here may serve as a conceptual framework for unifying the findings from studies that otherwise seemed unrelated (e.g., Degani et al., 2013; Fitzpatrick & Izura, 2011; Ko et al., 2011; Thierry & Wu, 2007) and for exploring issues such as how L2 speakers' L1, the learning history, and L2 proficiency may affect lexical representation and processing in an L2.

We want to end with two notes of caution. First, this is the first study where a translation frequency effect was reported. This finding has to be replicated. In replication, care should be given to two limitations pointed out by an anonymous reviewer: (a) the use of a more objective proficiency measure where proficiency is manipulated, and (b) increasing the statistical power of the design where necessary (see Brysbaert & Stevens, 2018). Second, while we interpreted the finding as evidence for verification, one may also consider if this translation frequency effect can be explained in alternative ways. For example, instead of proposing a strictly serial model in which verification occurs only after a single L2 word is identified, a cascaded processing model can also be conceived. In such a model, input generates multiple candidates before a single candidate was identified. Verification takes place

as soon as these candidates are generated rather than wait until a single candidate is identified. Thus, multiple candidates are verified. A model conceived this way would predict that processing the English word *teach* would not only activate the translation of this word but also the translation of *tea*. Further research is certainly needed to better understand the implications of the translation frequency effect for L2 word recognition.

Notes

1. See Thierry and Wu (2007) for explanations for the difference in the behavioral data.
2. See Wu and Thierry, 2012; Oppenheim, Wu, and Thierry, 2018; Wang, Wang, and Malins, 2017; Wen, Filik, and van Heuven, 2018, for evidence of L1 translation activation involving other paradigms.
3. Linguistically, the significantly higher frequency associated with the Chinese translations is mostly due to the fact they have a broader semantic coverage than their English counterpart. The word *yanjiu*, for example, covers all the meanings of its counterpart *research*, that is, to make a systematic investigation in an academic or scientific sense, but it is also used in situations where “to discuss” or “to talk about” are used in English. Culturally, some Chinese words are much higher in frequency than their English translations because of the particular political system or cultural practice of the Chinese society. For example, *tongxue* (classmate) represents a much more prominent personal relationship among Chinese than among Americans.
4. The RT data among the two NNS groups should not be interpreted as an indication of their English proficiency. Instead, age was likely to play a much more important role.

References

- Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory & Cognition*, **10**, 565–575.
- Athanasopoulos, P. (2007). Interaction between grammatical categories and cognition in bilinguals: The role of proficiency, cultural immersion, and language of instruction. *Language and Cognitive Processes*, **22**, 689–699.
- Athanasopoulos, P., Dering, B., Wiggett, A., Kuipers, J. R., & Thierry, G. (2010). Perceptual shift in bilingualism: Brain potentials reveal plasticity in pre-attentive colour perception. *Cognition*, **116**, 437–443.
- Becker, C. A. (1976). Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 556.
- Beijing Language Institute (BLI). (1986). *A frequency dictionary of modern Chinese*. Beijing: Beijing Language Institute Press.
- Brybaert, M., Lagrou, E., & Stevens, M. (2017). Visual word recognition in a second language: A test of the lexical entrenchment hypothesis with lexical decision times. *Bilingualism: Language and Cognition*, **20**, 530–548.
- Brybaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, **41**, 977–990.
- Brybaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, **1**, 1–20. doi:10.5334/joc.10
- Cheng, C., Wang, M., & Perfetti, C. A. (2011). Acquisition of compound words in Chinese–English bilingual children: Decomposition and cross-language activation. *Applied Psycholinguistics*, **32**, 583–600.
- Colombo, L., Pasini, M., & Balota, D. A. (2006). Dissociating the influence of familiarity and meaningfulness from word frequency in naming and lexical decision performance. *Memory & Cognition*, **34**, 1312–1324.
- Colomé, A. (2001). Lexical activation in bilinguals’ speech production: Language-specific or language-independent? *Journal of Memory and Language*, **45**, 721–736.

- Cook, V. J., Bassetti, B., Kasai, C., Sasaki, M., & Takahashi, J. A. (2006). Do bilinguals have different concepts? The case of shape and material in Japanese L2 users of English. *International Journal of Bilingualism*, *10*, 137–152.
- Costa, A., Pannunzi, M., Deco, G., & Pickering, M. J. (2017). Do bilinguals automatically activate their native language when they are not using it? *Cognitive Science*, *41*, 1629–1644.
- Degani, T., Prior, A., Eddington, C. M., Arêas da Luz Fontes, A., & Tokowicz, N. (2013). *Beyond semantic ambiguity: Effects of translation ambiguity in a single-language context*. Unpublished manuscript.
- de Groot, A. M. B., Borgwaldt, S., Bos, M., & Van den Eijnden, E. (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language*, *47*, 91–124.
- Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first- and second-language word recognition: A lexical entrenchment account. *Quarterly Journal of Experimental Psychology*, *66*, 843–863.
- Ecke, P. (2015). Parasitic vocabulary acquisition, cross-linguistic influence, and lexical retrieval in multilinguals. *Bilingualism: Language and Cognition*, *18*, 145–162.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116–124.
- Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp. 257–287). Amsterdam: North-Holland.
- Fitzpatrick, T., & Izura, C. (2011). Word association in L1 and L2. *Studies in Second Language Acquisition*, *33*, 373–398.
- Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, *33*, 1220–1234.
- Grainger, J. (1993). Visual word recognition in bilinguals. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 11–25). Amsterdam: Benjamins.
- Grainger, J., Midgley, K. J., & Holcomb, P. J. (2010). Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). In M. Kail & M. Hickman (Eds.), *Language acquisition across linguistic and cognitive systems* (pp. 267–284). Philadelphia: Benjamins.
- Green, D. W. (1998). Mental control of the bilingual lexicosemantic system. *Bilingualism: Language and Cognition*, *1*, 67–81.
- Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, *1*, 131–149.
- Hall, C. J., & Ecke, P. (2003). Parasitism as a default mechanism in L3 vocabulary acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *The multilingual lexicon* (pp. 71–85). Dordrecht: Kluwer Academic.
- Jared, D., & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language*, *44*, 2–31.
- Jiang, N. (1999). Testing processing explanations for the asymmetry in masked cross-language priming. *Bilingualism: Language and Cognition*, *2*, 59–75.
- Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics*, *21*, 47–77.
- Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language*, *49*, 484–499.
- Ko, I. Y., Wang, M., & Kim, S. Y. (2011). Bilingual reading of compound words. *Journal of Psycholinguist Research*, *40*, 49–73.
- Koda, K. (1989). Effects of L1 orthographic representation on L2 phonological coding strategies. *Journal of Psycholinguistic Research*, *18*, 201–222.
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest: Tests in linear mixed effects models. <https://CRAN.R-project.org/package=lmerTest>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*. doi:10.18637/jss.v082.i13
- Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38.

- Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*, 1494–1502.
- Meade, G., Midgley, K. J., Sehyr, Z. S., 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.
- Moon, J., & Jiang, N. (2012). Nonselective lexical access in different-script bilinguals. *Bilingualism: Language and Cognition*, *15*, 173–180.
- 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*, 286–292.
- Olshtain, E., & Cohen, A. (1989). Speech act behavior across languages. In H. W. Dechert & M. Raupach (Eds.), *Transfer in language production* (pp. 53–67). Westport, CT: Ablex.
- Oppenheim, G., Wu, Y. J., & Thierry, G. (2018). Found in translation: Late bilinguals do automatically activate their native language when they are not using it. *Cognitive Science*, *42*, 1700–1713.
- Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation–verification model for letter and word recognition: The word-superiority effect. *Psychological Review*, *89*, 573.
- Perea, M., & Carreiras, M. (1998). Effects of syllable frequency and syllable neighborhood frequency in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 134–144.
- Thierry, G., & Wu, Y. J. (2004). Electrophysiological evidence for language interference in late bilinguals. *NeuroReport*, *15*, 1555–1558.
- Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, *104*, 12530–12535.
- Tokowicz, N. (2014). Translation ambiguity affects language processing, learning, and representation. In R. T. Miller, K. I. Martin, C. M. Eddington, A. Henery, N. M. Miguel, A. M. Tseng, . . . Walter, D. (Eds.), *Selected Proceedings of the 2012 Second Language Research Forum* (pp. 170–180). Somerville, MA: Cascadilla Proceedings Project.
- van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, *39*, 458–483.
- Villameriel, S., Dias, P., Costello, B., & Carreiras, M. (2016). Cross-language and cross-modal activation in hearing bimodal bilinguals. *Journal of Memory and Language*, *87*, 59–70.
- Wade-Woolley, L. (1999). First language influences on second language word reading: All roads lead to Rome. *Language Learning*, *49*, 447–471.
- Wang, M., & Koda, K. (2005). Commonalities and differences in word identification skills among learners of English as a second language. *Language Learning*, *55*, 71–98.
- Wang, M., Lin, C. Y., & Gao, W. (2010). Bilingual compound processing: The effects of constituent frequency and semantic transparency. *Writing Systems Research*, *2*, 117–137.
- Wang, X., Wang, J., & Malins, J. G. (2017). Do you hear “feather” when listening to “rain”? Lexical tone activation during unconscious translation: Evidence from Mandarin-English bilinguals. *Cognition*, *169*, 15–24.
- Wen, Y., Filik, R., & van Heuven, W. J. B. (2018). Electrophysiological dynamics of Chinese phonology during visual word recognition in Chinese-English bilinguals. *Scientific Reports*, *8*, 68–69.
- Wen, Y., & van Heuven, W. J. B. (2017). Chinese translation norms for 1,429 English words. *Behavior Research Methods*, *49*, 1006–1019.
- Wen, Y., & van Heuven, W. J. B. (2018). Limitations of translation activation in masked priming: Behavioural evidence from Chinese-English bilinguals and computational modeling. *Journal of Memory and Language*, *101*, 84–96.
- Wu, Y. J., & Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *Journal of Neuroscience*, *30*, 7646–7651.
- Wu, Y. J., & Thierry, G. (2012). Unconscious translation during incidental foreign language processing. *NeuroImage*, *59*, 3468–3473.
- Zhang, T., van Heuven, W. J. B., & Conklin, K. (2011). Fast automatic translation and morphological decomposition in Chinese-English bilinguals. *Psychological Science*, *22*, 1237–1242.
- Zughoul, M. R. (1991). Lexical choice: Towards writing problematic word lists. *International Review of Applied Linguistics*, *29*, 45–59.

Appendix A Stimuli used in the two experiments

1. English targets of higher and lower translation frequency matched for lexical frequency

Higher translation frequency				Lower translation frequency					
Targets	F1	L	Chinese	F2	Targets	F1	L	Chinese	F2
revolution	12	10	革命	1963	signature	12	9	签名	3
society	33	7	社会	1841	contract	33	8	合同	34
develop	10	7	发展	1126	appetite	9	8	胃口	6
enemy	49	5	敌人	1078	punish	10	6	惩罚	15
research	33	8	研究	1001	faith	46	5	信仰	7
economy	6	7	经济	989	repeat	33	6	重复	49
chairman	12	8	主席	710	attract	6	7	吸引	51
study	49	5	学习	679	furniture	15	9	家具	30
history	84	7	历史	678	memory	49	6	记忆	62
politics	17	8	政治	631	evidence	85	8	证据	23
relation	4	8	关系	623	disaster	17	8	灾难	35
compare	15	7	比较	417	dictionary	4	10	词典	5
peasant	4	7	农民	290	apologize	8	9	道歉	12
content	8	7	内容	187	warning	32	7	警告	21
classmate	2	9	同学	153	liquid	8	6	液体	44
technology	15	10	技术	599	patience	15	8	耐心	48
<i>Mean</i>	<i>22.1</i>	<i>7.5</i>		<i>810.3</i>		<i>23.9</i>	<i>7.5</i>		<i>27.8</i>

2. English targets of high and low lexical frequency matched for translation frequency

High lexical frequency				Low lexical frequency					
guess	454	5	猜测	8	difficulty	6	10	困难	344
music	152	5	音乐	54	tuition	4	7	学费	11
room	440	4	房间	62	flood	6	5	水灾	4
husband	195	7	丈夫	62	hero	50	4	英雄	148
trouble	224	7	麻烦	63	allow	32	5	允许	102
understand	482	10	理解	77	suggest	39	7	建议	68
welcome	179	7	欢迎	109	pollution	2	9	污染	60
name	642	4	名字	138	wage	3	4	工资	49
morning	439	7	早晨	124	correct	45	6	纠正	56
remember	542	8	记得	148	carpet	12	6	地毯	17
season	187	6	季节	46	nurse	45	5	护士	43
church	136	6	教堂	13	battery	12	7	电池	18
explain	110	7	解释	86	scholar	4	7	学者	26
idea	359	4	想法	38	describe	17	8	描写	29
police	236	6	警察	81	accept	53	6	接受	126
company	147	7	公司	111	experiment	16	10	实验	144
<i>Mean</i>	<i>307.8</i>	<i>6.3</i>		<i>76.3</i>		<i>21.6</i>	<i>6.6</i>		<i>77.8</i>

Note: F1, English target frequency (per million). F2, frequency of the Chinese translations (per million). L, length (number of letters).

Cite this article: Jiang, N., Li, M., and Guo, T. (2020). A tale of two frequency effects: Toward a verification model of L2 word recognition. *Applied Psycholinguistics* 41, 215–236. <https://doi.org/10.1017/S0142716419000481>