THE JOURNAL OF ASIA TEFL
Vol. 19, No. 1, Spring, 2022, 50-65
http://dx.doi.org/10.18823/asiatefl.2022.19.1.4.50

The Journal of Asia TEFL

http://journal.asiatefl.org/

e-ISSN 2466-1511 © 2004 AsiaTEFL.org. All rights reserved.

Paired-Associate Second Language Vocabulary Learning: The Role of L1 Translation Familiarity

Hanzhong Sun
Quzhou University, China

Shaohua Fang
University of Pittsburgh, United States

First language (L1) use has always been a subject of debate in second language (L2) vocabulary learning, generating two main research lines. The first line concerns whether L1 use is advisable, and the second relates to how its use can be more beneficial. The current study contributes to the latter strand by probing the effect L1 familiarity had on L2 vocabulary gains using paired-associate learning paradigm. Two groups of English as a foreign language (EFL) learners with comparable L2 proficiency were recruited to learn three L2 word sets varying in L1 familiarity status (high/low), and accordingly three conditions were created (*High*: all L2 words with high L1 familiarity; *Low*: all low; *Mixed*: half high and half low). In the treatment session, both groups of participants studied each set separately, while made aware of, and then completing one of the two tests differing in depth of processing (either meaning recognition or meaning recall). At the testing phase, an unannounced one-week delayed posttest which measured meaning recognition was administered to both groups to examine the retention of vocabulary knowledge. The results showed that similar patterns were observed across two groups, such that *High* displayed an advantage over *Low* and *Mixed* did not exhibit striking differences compared to *High*. Additionally, recall group outperformed recognition group on all word sets. Therefore, our findings offer direct experimental support to the principle of learning burden (Nation, 1990, 2013), and highlight both the importance and the complexity of L1 familiarity. We conclude with pedagogical implications in relation to those findings.

Keywords: L1 familiarity effect, paired-associate learning, recognition and recall, form-meaning connection, depth of processing

Introduction

Paired-associate learning has long been acknowledged as a popular deliberate vocabulary learning paradigm in traditional classroom, where new L2 words are often presented (a) out of context and (b) with L1 translation equivalents attached behind. A quick glimpse of the EFL textbooks[1] and other learning materials at hand (e.g., vocabulary lists) revealed that most of them fit neatly with this practice, though a few also provide English synonyms or short explanations. Therefore, it comes as no shock to us that learners in this context rely heavily on their L1 to acquire L2 words as a shortcut strategy to rapidly expand their vocabulary size. Prior research indicated that this decontextualized translation learning was an effective method (Joyce, 2018; Sagarra & Alba, 2006; Webb, 2007) in terms of making form-meaning connections, and some studies even found that this context-free approach was equal or even superior to contextualized methods such as sentence learning (Mondria, 2003; Prince, 1996) and picture learning

(Lotto & de Groot, 1998) in acquiring form-meaning knowledge. Taken as a whole, these studies underscore the importance of applying L1 translation to enhance L2 vocabulary learning gains especially at the initial stage of learning. Given the mediation of L1 translation, its characteristics should be reasonably assumed to play a role in facilitating or constraining the outcomes.

As a response, a number of empirical studies have been undertaken to examine various L1 factors related to form, meaning and use. Some L1 factors are linked to form, such as cognateness (de Groot & Keijzer, 2000; Laufer & McLean, 2016; Urdaniz & Skoufaki, 2019); some are conceptually oriented, including L1-L2 congruence/overlap (Kweon, 2009; Wolter & Gyllstad, 2013; Wolter & Yamashita, 2015; Yamashita & Jiang, 2010; Zheng, 2011) and imaginability of concept (de Groot, 2006; de Groot & Keijzer, 2000; Ellis & Beaton, 1993); and a few are more contextual which comprise L1 frequency of occurrence (de Groot, 2006; de Groot & Keijzer, 2000; Lotto & de Groot, 1998; Paquot, 2017; Wolter & Gyllstad, 2013) and L1 familiarity (Chapman & Gilbert, 1937; Sun & Fang, 2021; Tagashira et al., 2010). Among these, research involving formal and conceptual factors has reached a consensus that words characterized as (a) cognate; (b) congruent between L1 and L2; (c) highly imaginable are easier to acquire/process. There is also no shortage of studies that investigated the effect of L1 frequency on L2 vocabulary use, learning and processing. Employing different tasks, these studies, however, have yielded inconsistent results with regard to L1 frequency. For example, studies examining L2 use reported a strong effect of L1 frequency (Paquot, 2017), and those adopting paired-associate learning format obtained a small effect (de Groot, 2006), while some L2 processing studies found no such evidence (Wolter & Gyllstad, 2013). The use of frequency data drawn from large and representative corpora comprising printed materials may sometimes be problematic, because they are unable to capture the speakers' varied language experience (Conklin, 2020). Instead, subjective familiarity ratings may be viewed as a more reliable index that aligns with day-to-day language exposure (ibid.). By comparison, L1 familiarity has clearly been under-researched, which motivates our current study.

In the present study, we attempted to inspect how L1 familiarity impacted L2 vocabulary gains operationalized as form-meaning connection making in paired-associate learning paradigm. This study sought to partially replicate and extend Tagashira et al.'s (2010) findings in such a manner that we used Chinese EFL learners as our participants and categorically related noun sets as learning materials, as opposed to Japanese-speaking counterparts and adjective antonym pairs in their original study. If the L1 familiarity effect emerged, we also aimed to explore whether depth of processing played a modulating role, which was manipulated by employing two learning tasks orienting learners' attention towards either the recognition or the recall of L1 translation. Hopefully, this endeavor is able to generate insights informing vocabulary pedagogical interventions.

Literature Review

L1 Versus L2 Use in L2 Vocabulary Learning

The use of L1 as an instructional tool has been well recognized in EFL classroom context. A case in point is the dominance of grammar-translation pedagogy in classroom teaching for quite a long period, though it was then heavily criticized and thus gradually lost its central place. L1 later regained its due attention since it was re-evaluated as a valuable resource that learners could exploit to enrich their learning (for these changes, see the survey article for more details, Hall & Cook, 2012). To date, it has therefore remained a subject of immense interest and controversy for language researchers as well as practitioners (Butzkamm, 2011; Cook, 2001; Cook, 2010; Shin et al., 2020; Tsagari & Giannikas, 2020).

To shed light on the relative efficacy of L1 versus L2, a substantial body of vocabulary studies were carried out to examine how glossing language impacted the learning of new words embedded in passages. For example, Laufer and Shmueli (1997) revealed the superiority of L1 glossing over L2 glossing. Miyasako (2002) found that the benefits of glossing language depended on learners' proficiency, such

that proficient learners gained more from L2 glossing while L1 glossing was more effective for their less proficient counterparts. These results were also confirmed in a recent meta-analytic research by Kim, Lee and Lee (2020). However, situated in the context of computer-assisted learning, Rouhi and Mohebbi's (2012) study revealed that learners' L1 worked as effectively as L2 glosses for vocabulary learning.

Aside from investigating L1 use in context, researchers also explored the advantages of using L1 translation to learn L2 words, a preferred option accessible to beginner learners, because textbooks and dictionaries intended for them normally afford L1 equivalents. In Lotto and de Groot's (1998) study, Dutch learners were instructed to learn unknown Italian words, either with their translation equivalents in Dutch (translation learning) or with a picture depicting their referents (picture learning) provided. Success in learning was defined as producing the correct Italian word in response to the picture or Dutch word (i.e., L1/picture-L2 translation). Translation learning was found to result in better performances compared to picture learning. Mondria (2003) contrasted the relative effectiveness of meaning-inferred method, where learners of French had to infer the meaning of unknown words from the context, and meaning-given method in which L1 translation was directly provided. The results indicated that both methods led to similar retention rate as measured by L1 translation recall, but meaning-given method consumed less time, thus more efficient. Joyce (2018) compared the learning efficiency of using L1 translations versus L2 definitions when Japanese learners of English studied academic vocabulary. Receptive knowledge tests revealed that receiving exposure to L1 translations contributed to larger learning gains.

Dissatisfied with the single aspect of lexical knowledge measured in previous studies, Webb's (2007) study tapped into various dimensions related to orthograph, association, grammatical function and form-meaning knowledge. One group of learners were asked to learn the target items presented together with their L1 equivalents. The other group completed a similar task where additional single sentence context for each word was also provided. He found that both tasks promoted large gains in all lexical dimensions under investigation, and no significant differences were demonstrated between these two groups.

Taken together, these studies consistently confirmed that L1 translation was a powerful resource learners could deploy to contribute to their L2 lexical acquisition, particularly form-meaning knowledge.

Role of L1 Translation Familiarity in L2 Vocabulary Learning

Considering the relative effectiveness of applying L1 translation to L2 vocabulary learning, attention has also been paid to how its use can reap maximal benefits. For example, researchers adopted paired-associate learning format to assess the role of L1 translation familiarity. Then why does L1 familiarity impact the learnability of L2 lexical items? Perhaps the most direct theoretical account connected with this is the general principle of learning burden postulated by Nation (2013).

> The more a word represents patterns and knowledge that the learners are already familiar with, the lighter its learning burden. These patterns and knowledge can come from the first language, from knowledge of other languages, and from previous knowledge of the second language. (emphasis added, pp. 44-45)

Viewed in a different light, familiar concepts are more deeply entrenched in the learners' mental lexicon, and because of their higher availability, corresponding L2 lexical forms can be more readily hooked onto these well-established concepts.

This theory has also been empirically corroborated (Chapman & Gilbert, 1937; Sun & Fang, 2021; Tagashira et al., 2010). For example, Chapman and Gilbert (1937) performed a word learning experiment in which English-speaking participants encountered unknown Hindustani words paired with English equivalents. Manipulations were conducted such that half of the English equivalents were familiar and the other half were unfamiliar to learners. A multiple-choice test demonstrated that English-Hindustani associations were formed more readily and permanently when English words (L1 translations) were of higher familiarity. By the same token, in Sun and Fang's (2021) study, Chinese EFL participants learned

English words with familiar and unfamiliar L1 equivalents in both semantically related and unrelated conditions. Translation recognition posttests showed that words with higher L1 familiarity were retained better in both conditions, indicating a strong positive L1 familiarity effect.

Differing from these two studies, Tagashira et al. (2010) varied the L1 familiarity across three types of L2 adjective antonym pairs. Specifically, both L2 words in the same pair received either high or low L1 familiarity ratings (High-High, *ample-meager;* Low-Low, *deductive-inductive*), or one word in a pair had high L1 familiarity and the other word was of low L1 familiarity (High-Low, *frail-robust*). This intentional manipulation intended to test the argument they proposed that similar L1 familiarity levels might cause additional interference. They conducted two studies where Japanese learners of English were instructed to perform a self-paced learning task. Multiple-choice tests were administered to explore how (a) high and low L1 familiarity; and (b) similar and different L1 familiarity levels influenced L2 antonym learning. When no time restrictions were imposed on learning session (Experiment 1), they found that on a one-week delayed posttest, Low-Low pairs were the most challenging, and High-High and High-Low pairs did not show any significant differences. However, in a follow-up study (Experiment 2), they replicated their Experiment 1 with two modifications. Time constraint was placed on learning phase and the same delayed posttest was administered three weeks after learning, as opposed to one week in Experiment 1. Interestingly, they found that High-High and Low-Low pairs displayed similar learning rates, both of which were worse than High-Low pairs. These results apparently contradicted the learning burden principle (Nation, 1990, 2013), since L1 familiarity effect did not emerged. They interpreted it as the result of interference caused by the similar familiarity (High-High) being larger than the superior effect contributed by higher familiarity. These findings, on the other hand, lent support to their proposed claim regarding the interference brought about by similar L1 familiarity levels.

Overall, these experiments were set up to investigate both the quantitative (high versus low) and the qualitative (similar versus different) differences of L1 familiarity. The findings derived from these studies have deepened our understanding about the role of L1 familiarity in L2 vocabulary learning. On a practical level, they also informed us of how new words should be organized as regards the employment of L1 translation so that the highest learning efficiency could be achieved. Despite the usefulness of the insightful implications, these earlier inquiries are not without limitations, especially Tagashira et al.'s (2010) study. For example, although their participants were required to study the word pairs carefully until they could provide correct L1 translations for all L2 words, no immediate posttest was administered to ensure this, or at least the authors did not explicitly specify this. It might be difficult to determine whether the mean differences in lexical performances on the delayed posttest were all ascribed to the treatment in the study. Additionally, since they solely focused on L2 adjective antonym word pairs, it remains to be seen whether their findings are applicable to words (a) of other class, such as noun; (b) presented in a word set, a larger unit than a word pair; (c) categorically linked to each other.

Inspired by our review of existing studies and identified shortcomings, the present study sought to investigate the effect of L1 familiarity on paired-associate L2 word learning by targeting Chinese learners of English and using a different set of materials, thus replicating and extending the results of Experiment 1 carried out by Tagashira et al. (2010). We administered two posttests in the hope of gaining a fuller picture concerning how L1 familiarity affected lexical performances. The immediate posttest aimed to ensure that the learners mastered the word-meaning knowledge as required, and the use of one-week delayed surprise posttest was to assess the retention of the lexical knowledge.

Recall and Recognition in Vocabulary Learning

In vocabulary studies, recall and recognition are the most widely employed test format. In a meaning recall test, for instance, participants are instructed to provide an explanation for the target item with no option available. A meaning recognition test, on the other hand, requires them to choose from an option pool the most appropriate meaning based on a given word. Prevailing in memory research, this pair of terms is often used interchangeably with *production* and *reception* in language learning literature.

Generally, recalling an item is intuitively more challenging than simply recognizing it, as recall entails more lexical information such as orthography.

Several rounds of hot theoretical inquiries were also launched into their respective psychological processes (Kintsch 1970; Tulving, 1974; Tversky, 1973). Among these discussions, two-stage theory posits that recall and recognition are two qualitatively distinct processes in that the former entails both retrieval and discrimination, whereas the latter only involves the discrimination stage[1] (Kintsch, 1970). In other words, recall requires deeper level of processing, thus demanding more cognitive efforts. Because of this, it is anticipated that participants will likely secure better performances when learning for recall than for recognition. Conversely, the continuity theory claims that they constitute two similar processes, and the difference resides in whether the retrieved information is presented together with the test items (Tulving, 1976). Either way, Tversky (1973) argued that learners might encode the item information in a different fashion when anticipating a recall versus recognition test, due in part to the fact that the correct alternatives are presented in a recognition test. To make it more accessible, she drew an analogy between language learning and making the acquaintance of a new person. In the latter scenario, participants tended to employ different encoding strategies when told to recall the name than to recognize the face.

In our study, we were also interested in finding out whether the emergence of L1 familiarity effect was affected by depth of processing during treatment session, represented by two tasks where our participants were asked to learn the words in anticipation of either a L1 recall or a L1 recognition test.

Aims and Research Questions

This study represented an endeavor to explore the extent to which L1 translation familiarity impacted L2 lexical performances (i.e., form-meaning recognition) adopting paired-associate learning paradigm. Following Tagashira et al. (2010), we created three word sets, and manipulated two dimensions regrading L1 familiarity, i.e., high/low and similar/different. Comparisons across these three word sets were conducted to reveal the effect of varying L1 familiarity. The second aim was to examine whether depth of processing modulated the emergence of this effect. To do this, one group of learners performed a word learning task while made fully aware that there would be an L1 recall test, so that during the learning session they should be able to retrieve L1 equivalents in response to L2 words; learners in the other group was involved in the same task, but they completed a recognition test with L1 equivalents given. Based on these descriptions, three specific research questions (RQ) were addressed.

RQ1: Do high L1 familiarity ratings (as opposed to low) lead to better L2 lexical performances?
RQ2: Do different L1 familiarity ratings (as opposed to similar) lead to better L2 lexical performances?
RQ3: Does the effect of L1 familiarity, if detected, depend on depth of processing represented by recognition and recall tasks during learning?

Method

Participants

The present study was implemented in a university in mainland China. Through convenient sampling, a total of 60 participants were recruited from two intact classes (labeled as Group A and Group B respectively) where the first author taught. All the participants were Chinese L1 speakers and enrolled as

[1] The later version of this theory was revised, stating that the recognition process also includes retrieval (Tulving & Thomson, 1973).

full-time freshman engineering undergraduates. 14 participants did not complete the study as required, due to being absent from class and leaving part of information unfilled or unidentified. They failed to meet the inclusion criteria, thus not included in further analysis. Biographical data for the remaining 46 participants regarding their age, years spent on English learning and English proficiency were calculated and displayed in Table 1. It is important to note that we did not use a proficiency test to assess their English ability due to practical constraints, but following Sonbul and El-Dakhs's (2020) study, we employed the updated Vocabulary Levels Test (Updated VLT; Webb et al., 2017) as a proxy for their general English language proficiency. Specifically, the participants were measured on the first three levels (1000, 2000, 3000) of the Updated VLT (Version B) and the total score was 90.

We ran two independent samples t-tests to determine whether the participants in two groups had comparable length of time spent on learning English and English proficiency. The results provided affirmative answers to both. The participants in both groups learned English for approximately 10 years on average (SD_A = 1.04, SD_B = 0.85), and the between-group differences were not statistically significant (t (44) = 0.93, p = .356). For the vocabulary test, the mean total score for each group was about 80 points (SD_A = 5.46, SD_B = 5.28), and the differences between these two groups also did not reach significance level (t (44) = -0.14, p = .891). Additionally, both groups of participants received English instructions mainly in classroom setting, and vocabulary lists featured by L2-L1 pairs were regarded as the main channel through which they acquired L2 words, although other sources were identified such as watching movies. It should be therefore the case that Group A and Group B were homogeneous in terms of both English learning experience and English proficiency, and paired-associate learning was considered as a very natural method they probably got accustomed to.

TABLE 1
Biographical Data of the Participants

	N	Age	Years Learning English	Updated VLT
Group A	23	19.13 (0.76)	10.48 (1.04)	80.52 (5.46)
Group B	23	19.39 (0.50)	10.22 (0.85)	80.73 (5.28)

Item Development

Following paired-associate learning paradigm, the target items used in this study were three sets of novel English words presented with their respective L1 (Chinese) equivalents. Words in each set were of the same category. We manipulated the familiarity status of these Chinese equivalents, creating three conditions: High, Low and Mixed. *High* refers to the set within which the L1 equivalents of all L2 words were familiar to our participants; *Low* means that L1 equivalents of these words were all unfamiliar to the participants. In *Mixed* set, we included half of the L2 words whose L1 equivalents were familiar while the other half of the L2 words whose L1 counterparts were not.

We started with a pool of candidate Chinese words from three broad categories (i.e., vegetable, musical instrument and bird). This selection was based on our intuitive judgement of their visibility or availability in real life. For instance, vegetables are indispensable part of our daily consumption making most of them conceptually familiar to us, whereas musical instruments are not. We consulted various resources (*Oxford Advanced Learner's Dictionary*, 2020; Sun & Fang, 2021) from which 20 candidate Chinese words for each category were identified. We then recruited another batch of 42 learners sampled from the same population but not attending the main experiment. They were asked to rate the familiarity of these 60 Chinese words on a 7-point scale, with "1"signifying the least familiarity and "7" indicating the highest degree of familiarity. Based on the ratings, we included in each set 10 Chinese words. All of them are composed of two Chinese characters, except for one (*响葫芦*) in musical instrument set. The mean ratings are presented in Table 2, and the complete word list is attached in Appendix. We further divided Mixed set into Mixed_High and Mixed_Low to demonstrate the differences more clearly. A one-way ANOVA test was performed to check the appropriateness of word selection. The results indicated significant differences in familiarity ratings among these sets (F(3,26) = 440.05, p < .0001). We then

conducted an array of LSD multiple comparison tests to reveal the differences. We did not observe any differences between High and Mixed_High (p = .342), and between Low and Mixed_Low (p = .550), but notable differences emerged between High and Low (p < .001), High and Mixed_Low (p < .001), Mixed_High and Low (p < .001), Mixed_High and Mixed_Low (p < .001). Therefore, the selected words met our requirements.

After finalizing the Chinese words, we continued to pair each of them with an English counterpart. Of note, we used English pseudowords exclusively for two purposes. The first purpose was to ensure that our participants had no pre-existing knowledge regarding these words, so we would not bother with administering a pre-test. The second one was to prevent autonomous learning between the learning and delayed posttest session, since they would not encounter these words in other materials outside of class. To create these novel words, we followed two steps. First, we picked from self-made vocabulary lists 30 real English words which were of low frequency. All these words were submitted to *Corpus of Contemporary American English*, producing very few occurrences, roughly 10 to 1,300 per billion words. They ranged from 5 to 7 in the number of letters, and from 2 to 3 in the number of syllables. Then, we changed one letter of each word, either a consonant or a vowel. All the changed novel words remained phonologically plausible, meaning that the participants would not be aware of the lexical status of these words if not told so.

With all these steps completed, each of the Chinese words (L1 translation) were randomly assigned with a novel English (L2) word. Because word length has been found to affect visual word learning (Ellis & Beaton, 1983), we ran another one-way ANOVA test to confirm whether L2 length was equivalent across the three word sets. The descriptive statistics was summarized in Table 3. The results revealed no differences in the number of letters (F(2,27) = 0, p = 1.000) and in the number of syllables (F(2,27) = 0.155, p = .857).

TABLE 2
Descriptive Statistics of Familiarity Ratings of Chinese Words

Word set	Category	N	Mean	*SD*	Min	Max	Example words
High	Vegetable	10	6.21	0.50	5.24	6.79	韭菜、山药
Low	Musical instrument	10	1.91	0.21	1.50	2.19	扁琴、黑管
Mixed_High	Bird	5	6.39	0.11	6.19	6.50	麻雀、喜鹊
Mixed_Low	Bird	5	1.80	0.21	1.55	2.00	山鹑、苍鹭

TABLE 3
Descriptive Statistics of Novel English Words

Word set	Category	N	No. of letters	No. of syllables	Example words
High	Vegetable	10	6.10 (0.88)	2.10 (0.32)	dulket, aspew
Low	Musical instrument	10	6.10 (0.57)	2.00 (0.00)	kismet, vipen
Mixed	Bird	10	6.10 (0.88)	2.10 (0.74)	smulk, jacose

Instruments

We designed three vocabulary tests in relation to the target words, with two taking the form of meaning recognition test (Version A and Version B) and one as meaning recall test. In this suite of instruments, we also included one difficulty ranking questionnaire, one validated vocabulary levels test and one background questionnaire, in an effort to obtain a more comprehensive picture of the participants.

Meaning recognition test (Version A) was administered during treatment session for one group of participants to elicit the form-meaning knowledge. In this test, ten English target words which the participants were instructed to memorize were given, together with their corresponding Chinese equivalents (options). They were asked to discriminate and match English (L2) words with Chinese (L1) translations. As well informed of the existence of this testing format, they were therefore encouraged to learn from L2 to L1 (e.g., *aspew* = ??). This L2-L1 direction (rather than L1-L2 direction) was preferable and recommended especially in EFL classroom context, because it was more effective for lower-

proficiency learners and less affected by proficiency (Terai et al., 2021). The construction of the following two vocabulary tests were also based on this consideration.

Meaning recall test was also used during treatment session, but intended for another group of participants. Unlike the recognition test, recall test only had English words presented with no provision of any Chinese equivalents for them to choose from. They were required to retrieve the Chinese equivalents from their memory.

Meaning recognition test (Version B) was identical to that of Version A, except for the sequencing of the items and options. This test served as the delayed test administered to both groups one week after the treatment session, and the participants were completely unaware of its presence. This one-week delayed posttest made us understand how varying L1 familiarity might influence retention. We also would like to point that using the same meaning recognition test here allowed direct comparisons across these two groups. The emerged patterns could inform how depth of processing during learning played a role.

Difficulty ranking questionnaire aimed to elicit the perception of the difficulty level of each set, thus triangulating the experimental data. The participants were asked to rank the three sets in ascending order, and specify the rationales for their ranking.

Updated Vocabulary Levels Test is a validated test that measures receptive vocabulary knowledge developed by Webb et al. (2017). It was used as a rough estimate for general English proficiency in the current study. It exists in two forms. In each form, there are five frequency bands of words ranging from 1,000 to 5,000. For each band, 30 definitions and 60 words (30 keys and 30 distractors) are presented. The test takers were required to match each word with the corresponding definition. Due to the time limits, we only used the first three bands.

Background questionnaire was deployed to collect information related to the participants' English learning, such as age, years spent on English learning and the resource they have been relying on to learn new words (they were told to indicate the main one if multiple resources were identified).

Procedure

This experiment was performed in two intact classes spanning three weeks. In the first week, we secured informed consent in an oral form from the participants and briefed them on the purpose of this research, i.e., to see how well they could commit given words to memory within ten minutes[2]. They were then asked to complete one background questionnaire and one vocabulary test.

In the second week, the two classes were randomly assigned to either recognition or recall treatment session. During recognition treatment session, they were handed out a word card containing the target items and instructed to learn these words at their own pace, and meanwhile made totally aware that there would be a vocabulary test (meaning recognition test, version A) immediately after learning, where they were asked to match the English words with a pool of Chinese words available. Learners receiving recall treatment session were presented with the same list of words, and performed the memorizing task also with the upcoming vocabulary test (meaning recall test) bearing in mind. The only difference was that there were no Chinese words offered as options in this test, so they had to recall from memory. Put succinctly, both groups of students were required to use English (L2) words as stimuli and Chinese (L1) counterparts as responses to create all the links between L2 form and L1 translation during treatment, while anticipating either L1 recognition or L1 recall test. It should be mentioned that prior to the test, the word card was collected to prevent further learning. One week after learning (the third week), both groups re-took a vocabulary test (meaning recognition test, version B) with no prior announcement. Both groups of participants were engaged in learning and testing one set of words at a time and they therefore all underwent three rounds. Because of this, learning order might constitute another confounding factor. To avoid possible effects arising out of it (i.e., recency effect), both groups were further separated into three subgroups. Within each group, a counterbalanced design was adopted, which was clearly demonstrated in Table 4. Approaching the end of this experiment, learners were also asked to rank the difficulty level of each set and explain concisely. Upon completing all sessions, the participants were informed of the real

purposes of this study, and the real word lists were given to them so that they would not use the pseudowords in their future production.

The scoring of the tests was done by the first author. One point was awarded to one correct response. Wrong or null responses were not given any point. The total score of each test was 10.

TABLE 4
The Procedure of the Experiment

Week	Day	Procedure					
1	1	Background questionnaire Updated Vocabulary Levels Test					
		Recognition Group			Recall Group		
		Subgroup 1	Subgroup 2	Subgroup 3	Subgroup 1	Subgroup 2	Subgroup 3
2	1	High	Low	Mixed	High	Low	Mixed
(Learning	2	Low	Mixed	High	Low	Mixed	High
+ IT)	3	Mixed	High	Low	Mixed	High	Low
3	1	High	Low	Mixed	High	Low	Mixed
(DT)	2	Low	Mixed	High	Low	Mixed	High
	3	Mixed	High	Low	Mixed	High	Low
	4	Difficulty ranking questionnaire					

Note. IT= immediate posttest, DT= one-week delayed posttest

Results

We calculated the scores of each word set for both groups on the two tests. The immediate posttest showed that two groups of participants did magnificently well by achieving 100% accuracy rate in each set, indicating that they mastered all the words at the initial stage of learning in terms of form-meaning connection.

However, differences emerged on the delayed posttest. Its descriptive statistics was presented in Table 5 and Table 6. The one-way ANOVA was conducted separately for each group to analyze whether the differences in lexical performances resulted from L1 familiarity. Within the recognition group, we found a strong L1 familiarity effect ($F(2,66) = 7.93$, $p = .001$). The post-hoc Tukey's HSD tests were then carried out to reveal how each set differed at statistical significance level, which was set at 0.05. The results demonstrated that the differences between High and Low ($p = .018$, 95%CI = [0.27, 3.47]), and between Mixed and Low ($p = .001$, 95%CI = [0.97, 4.16]) were significant, but the differences between High and Mixed ($p = .522$, 95%CI = [-2.29, 0.90]) were not. Similar statistical tests were performed for the recall group. We also obtained a strong effect of L1 familiarity ($F(2,66) = 5.66$, $p = .005$). Subsequent Tukey's HSD multiple comparison tests showed that significant differences were detected between High and Low ($p = .019$, 95%CI = [0.21, 2.83]), and between Mixed and Low ($p = .01$, 95%CI = [0.34, 2.96]), but not present between High and Mixed (p = .969, 95%CI = [-1.44, 1.18]). We also ran three independent-samples t-tests to how the two groups differed from each other in three word sets. The results revealed that recognition group outperformed recall group on all of these word sets ($p < .001$). Taken together, our results exhibited that both L1 familiarity and depth of processing exerted a strong influence on L2 word learning.

These findings above therefore provided direct answers to the three research questions addressed in this study. L1 familiarity was positively associated with L2 lexical performances, but it did not necessarily result in higher L2 lexical gains, because High and Mixed did not exhibit differences whereas Mixed and Low did. Since the same pattern was found in both groups, the emergence of L1 familiarity effect was not dependent upon depth of processing.

TABLE 5

Descriptive Statistics of the Performances for Recognition Group on the Delayed Posttest

	N	Mean	*SD*	Std. Error	Min	Max	95%CI
High	23	4.52	2.41	0.50	1	10	[3.48, 5.56]
Low	23	2.65	1.94	0.41	0	7	[1.81, 3.49]
Mixed	23	5.22	2.39	0.50	2	10	[4.18, 6.25]

Note. CI = confidence interval

TABLE 6
Descriptive Statistics of the Performances for Recall Group on the Delayed Posttest

	N	Mean	*SD*	Std. Error	Min	Max	95%CI
High	23	9.48	0.95	0.20	7	10	[9.07, 9.89]
Low	23	7.96	2.88	0.60	2	10	[6.71, 9.20]
Mixed	23	9.61	1.03	0.22	6	10	[9.16, 10.06]

Aside from the test results, we also distributed a questionnaire at the end of the experiment to collect their responses to the perception of the ease or difficulty associated with each set. The ranking data were reported in Table 7. It should be noted that the results were not quite clear-cut, because responses regarding this tended to be highly subjective. However, we could identify certain patterns from these messy data points. For example, words of high L1 familiarity were more likely to be rated as *easy* to learn (56.53% and 69.57%, respectively), and those of low L1 familiarity was the least likely to be done so (8.70% and 8.70%). Learners in recognition group (34.78%) were more inclined to view the mixed set as *easy* than those in recall group (13.03%). Overall speaking, the subjective difficulty rankings on High and Low aligned closely with their actual performances on the delayed posttest. This was true for both groups. Their rankings on Mixed displayed a wide divergence from the test scores. We also sought the rationales behind the rankings, and found that the (un)familiarity with Chinese equivalents was the main consideration when making the ranking decisions.

TABLE 7
Self-Perception of the Difficulty Level towards Each Word Set

	Recognition group			Recall group		
	1(easy)	2	3 (difficult)	1(easy)	2	3 (difficult)
High	13 (56.53%)	7 (30.43%)	3 (13.04%)	16 (69.57%)	6 (26.09%)	1 (4.35%)
Low	2 (8.70%)	10 (43.48%)	11(47.83%)	2 (8.70%)	9 (39.13%)	12 (52.17%)
Mixed	8 (34.78%)	6 (26.09%)	9 (39.13%)	3 (13.03%)	12 (52.17%)	8 (34.78%)

Discussion

The current study intended to probe into how varying degrees of L1 familiarity (i.e., High, Low and Mixed) impacted the learnability of L2 words. Paired-associate learning format was adopted where participants were instructed to associate L2 words with their Chinese translations while expecting either L1 recognition or L1 recall test immediately after learning. Our results showed that the participants were successful in acquiring the connections between L2 form and L1 meaning, which gives clear evidence of the effectiveness of this method. It might be due to the fact that learners were given ample time and informed of the existence of the test. This finding is also consistent with previous studies (Joyce, 2018; Sagarra & Alba, 2006; Webb, 2007). To assess how well these connections were retained, a subsequent delayed posttest measuring L1 recognition was deployed. It is worth mentioning that because our learners achieved 100% accuracy on the immediate posttest, any difference in the performances on the delayed posttest could be attributed to treatment effect (i.e., L1 familiarity). A few interesting patterns shared by both groups were obtained. In what follows, we elaborated on each finding in relation to the research questions.

The first research question asked whether higher L1 familiarity contributed to better L2 vocabulary learning at the early stage. Performances of high and low L1 familiarity words (i.e., High and Low) were compared. The results exhibited that words with higher L1 familiarity were easier to retain in the mental lexicon than those whose L1 equivalents were less familiar. This was indicative of the strong role L1 familiarity played in L2 lexical acquisition. Our results overlap the findings of earlier studies (Chapman & Gilbert, 1937; Sun & Fang, 2021; Experiment 1 of Tagashira et al., 2010). These studies, including the current one, unequivocally confirmed the presence of L1 familiarity effect, though participants from a variety of L1 backgrounds were taken as samples (i.e., English, Chinese and Japanese). Our finding elicited from the difficulty ranking questionnaire also lent support to the test performances, indicating that words receiving high L1 familiarity ratings tended to be considered as relatively easy to learn. Informed by the principle of learning burden (Nation, 1990, 2013), one plausible reason for the advantage of high over low L1 familiarity might be that familiar L1 concepts were already entrenched in the lexical network and therefore more ready to be deployed to develop L2-L1 associations. As a result, learners were likely to experience less cognitive burden on acquiring this type of lexical items. On the other hand, when encountered with words of low L1 familiarity, they probably struggled with getting familiarized with the concepts first, which might take huge amounts of mental efforts. Even though L2-L1 connections were successfully created during learning session, they might be weak and unstable because L1 concepts which L2 forms hooked onto were not deeply rooted in learners' mental lexicon.

The second research question investigated whether clustering words of different L1 familiarity ratings led to better L2 performances. To do this, we compared words that received similar versus different L1 familiarity ratings (i.e., High and Mixed, Low and Mixed). Our results indicated that no differences emerged between High and Mixed. This seemed to conflict with the positive L1 familiarity effect addressed in the preceding text, because words in High (M = 6.21) were generally rated as more familiar in comparison with those in Mixed (M = 4.10). To account for the absence of L1 familiarity effect, we concurred with Tagashira et al.'s (2010) proposal that presenting semantically related words with similar L1 familiarity might introduce additional interference. In the current study, when learners were exposed to words in High, they benefited from these highly familiar L1 concepts so that L2-L1 links were developed rapidly, but meanwhile these L1 concepts competed against each other, which led to confusion over time. As a consequence, L1 familiarity effect might be offset by this interference effect making it unable to be observed. However, we did find meaningful differences between Low and Mixed. Words in Mixed outperformed their counterparts in Low. One potential reason why the set containing words of both high and low L1 familiarity were advantageous, as opposed to the set with low L1 familiarity words only, might lie in the L1 familiarity effect. In the mixed set, an average L1 familiarity rating of 4.10 was obtained, whereas the low set only received a rating of 1.91. Another possible reason why the low set was disadvantaged might be the competition caused by similar L1 familiarity levels. Taken together, we successfully replicated the first experiment of Tagashira et al. (2010) using Chinese EFL learners with similar results, and extended theirs from L2 antonym pairs to L2 categorically related word sets.

The third research question was concerned about the extent to which depth of processing represented by two tasks (i.e., recognition and recall) during learning session affected the L1 familiarity effect. One group of participants performed the word learning task in order to complete an L1 recognition test, so they did not bother to memorize all L1 equivalents. The other group learned the same word sets but completed a test that required them to retrieve all L1 translations. Though the completion time of learning was not precisely recorded, it took overall shorter time for the recognition group. Our data displayed the same patterns for these two groups concerning how L1 familiarity affected L2 performances, and therefore we argue that L1 familiarity, as an inherent intra-lexical property, was independent of processing depth. However, striking differences could be observed across the two groups in terms of learning gains. We interpret these different results as the evidence of the depth of encoding during learning session. Leaners in recall group needed to fully encode the L1s in their mental lexicon for subsequent successful retrieval, but partial encoding would suffice for the recognition group in a manner where learners were able to discriminate different L1s. Pedagogically, this also conforms to one of the

basic tenets in second language acquisition that learners need to be challenged to learn. These findings, nonetheless, put us in no position to argue whether recognition and recall are qualitatively similar or different processes.

Conclusion

This study focused on L1 translation familiarity and investigated how its use could maximize the retention of L2 vocabulary that was categorically related. Specifically, we used three word sets (i.e., High, Low and Mixed) to assess how high/low and similar/different L1 familiarity impacted L2 vocabulary gains. Two groups of learners with comparable L2 proficiency conducted a timed but self-paced word learning session for subsequent either L1 recognition or L1 recall test. The results of one-week delayed posttest measuring L1 recognition knowledge indicated that (a) L1 familiarity was positively linked with L2 lexical performances; (b) different L1 similarity led to better L2 vocabulary gains but it was not always the case, suggesting the intricacy of L1 familiarity; and (c) learners in recall group excelled those of the recognition group in all three word sets, but both groups exhibited the same patterns as to how L1 familiarity affected L2 vocabulary learning. These findings support the principle of learning burden (Nation, 1990, 2013), and highlight the importance of L1 familiarity in creating pedagogical vocabulary activities.

This classroom-based study has several important implications for acquiring semantically related L2 words at the initial stage of learning. First of all, learners benefit from familiar L1 concepts which serve as an effective cognitive hook which L2 forms later hang onto (Fraser, 1999), and therefore it is worthwhile to get familiarized with L1 referents first. Furthermore, semantically related words with high L1 familiarity are not necessarily easy to acquire, but rather the learnability depends on what kind of L1 translations are clustered together (Tagashira et al., 2010). For instance, low L1 familiarity words are strongly advised to be presented with high L1 familiarity ones, as our research showed that lexical items with both low and high L1 familiarity were acquired as well as the items whose L1 translations were all familiar. Finally, recalling activities are recommended as a gateway to consolidation, as they require more retrieval efforts and lead to better learning than recognition exercises.

Despite the aforementioned valuable suggestions, a few limitations should be noted. First, the current study tapped into form-meaning recognition knowledge, so whether the results can be generalized to other components of vocabulary knowledge remains largely unknown. Second, we recruited two groups of participants, and each group were supposed to employ different encoding strategies (i.e., partial and full encoding of L1). However, no measure was taken to ensure this. Some learners in the recognition group might also memorize all the L1 equivalents. Future research may also profit by including a few learning- and learner-related factors, such as learning direction (L1 to L2 and L2 to L1, Terai et al., 2021) and pre-existing vocabulary knowledge (Webb & Chang, 2015).

Notes

1. We examined the English textbooks specifically written for non-English major undergraduates (our participants) and published by various prestigious presses in mainland China, including Shanghai Foreign Language Education Press, Foreign Language Teaching and Research Press, and Higher Education Press.
2. We ran a small-scale pilot study where three students who did not attend the main experiment were asked to learn one word set by rote. We found that ten minutes was a proper time frame.

The Authors

Hanzhong Sun (first author) is currently a lecturer at College English Department, College of Foreign Languages and International Education, Quzhou University, China. His primary research interest lies in (instructed) second language acquisition. Along this strand, he is particularly interested in such topics as L2 vocabulary acquisition, written corrective feedback and teacher talk. He also takes a profound interest in second language processing, especially L2 lexical processing and various lexical features associated with this process.

College English Department
College of Foreign Languages and International Education
Quzhou University
Quzhou, China, 324000
TEL: + 86 0570-8026507
Email: sunhanzhong@foxmail.com

Shaohua Fang (co-author) is currently a PhD student in linguistics at the University of Pittsburgh. His research interest lies in second language acquisition from the perspectives of theoretical linguistics and psycholinguistics. He is particularly interested in examining the differences between native and non-native speakers in the processing mechanisms involved.

Department of Linguistics
Kenneth P. Dietrich School of Arts and Sciences
University of Pittsburgh
Pittsburgh, 15260, United States
MOBILE: + 001 631-428-9203
Email: shf64@pitt.edu

References

Butzkamm, W. (2011). Why make them crawl if they can walk? Teaching with mother tongue support. *RELC Journal*, *42*(3), 379-391. https://doi.org/10.1177/0033688211419830

Chapman, F. L., & Gilbert, L. C. (1937). A study of the influence of familiarity with English words upon the learning of their foreign language equivalents. *Journal of Educational Psychology*, *28*(8), 621-628. https://doi.org/10.1037/h0063528

Conklin, K. (2020). Processing single-word and multiword items. In Stuart Webb (Eds.), *The routledge handbook of vocabulary studies* (pp. 174-188). Routledge.

Cook, G. (2010). *Translation in language teaching: An argument for reassessment*. Oxford University Press.

Cook, V. (2001). Using the first language in the classroom. *Canadian Modern Language Review*, *57*(3), 402-423. https://doi.org/10.3138/cmlr.57.3.402

de Groot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, *50*(1), 1-56. https://doi.org/10.1111/0023-8333.00110

de Groot, A. M. B. (2006). Effects of stimulus characteristics and background music on foreign language vocabulary learning and forgetting. *Language Learning*, *56*(3), 463-506. https://doi.org/10.1111/j.1467-9922.2006.00374.x

Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, *43*(4), 559-617. https://doi.org/10.1111/j.1467-1770.1993.tb00627.x

Fraser, C. A., (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition*, *21*(2), 225-241. https://doi.org/10.1017/s0272263199002041

Hall, G., & Cook, G. (2012). Own-language use in language teaching and learning. *Language Teaching*, *45*(3), 271-308. https://doi.org/10.1017/s0261444812000067

Joyce, P. (2018). L2 vocabulary learning and testing: The use of L1 translation versus L2 definition. *The Language Learning Journal*, *46*(3), 217-227. https://doi.org/10.1080/09571736.2015.1028088

Kim, H. S., Lee, J. H., & Lee, H. (2020). The relative effects of L1 and L2 glosses on L2 learning: A meta-analysis. *Language Teaching Research*. https://doi.org/10.1177/1362168820981394

Kintsch, W. (1970). Models for free recall and recognition. In D.A. Norman (Ed.), *Models of human memory* (pp. 331-373). Academic Press.

Kweon, S. (2009). Differential difficulty in processing L1-L2 verb splits: Dissociation with corpus frequency. *The Journal of Asia TEFL*, *6*(3), 1-25.

Laufer, B., & McLean, S. (2016). Loanwords and vocabulary size test scores: A case of different estimates for different L1 learners. *Language Assessment Quarterly*, *13*(3), 202-217. https://doi.org/10.1080/15434303.2016.1210611

Laufer, B., & Shmueli, K. (1997). Memorizing new words: Does teaching have anything to do with it? *RELC Journal*, *28*(1), 89-108. https://doi.org/10.1177/003368829702800106

Lotto, L., & de Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning*, *48*(1), 31-69. https://doi.org/10.1111/1467-9922.00032

Miyasako, N. (2002). Does text-glossing have any effects on incidental vocabulary learning through reading for Japanese senior high school students? *Language Education & Technology*, *39*, 1-20. https://doi.org/10.24539/let.39.0_1

Mondria, J. A. (2003). The effects of inferring, verifying, and memorizing on the retention of L2 word meanings: An experimental comparison of the "meaning-inferred method" and the "meaning-given method". *Studies in Second Language Acquisition*, *25*(4), 473-499. https://doi.org/10.1017/s0272263103000202

Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Newbury House.

Nation, I. S. P. (2013). *Learning Vocabulary in Another Language(2nd Ed)*. Cambridge University Press.

Oxford Advanced Learner's Dictionary(10th Ed). (2020). Oxford University Press.

Paquot, M. (2017). L1 frequency in foreign language acquisition: Recurrent word combinations in French and Spanish EFL learner writing. *Second Language Research*, *33*(1), 13-32. https://doi.org/10.1177/0267658315620265

Prince, P. (1996). Second language vocabulary learning: The role of context versus translations as a function of proficiency. *The Modern Language Journal*, *80*(4), 478-493. https://doi.org/10.1111/j.1540-4781.1996.tb05468.x

Rouhi, A., & Mohebbi, H. (2012). The effect of computer assisted L1 and L2 glosses on L2 vocabulary learning. *The Journal of Asia TEFL*, *9*(2), 1-19.

Sagarra, N., & Alba, M. (2006). The key is in the keyword: L2 vocabulary learning methods with beginning learners of Spanish. *The Modern Language Journal*, *90*(2), 228-243. https://doi.org/10.1111/j.1540-4781.2006.00394.x

Shin, J., Dixon, L., & Choi, Y. (2020). An updated review on use of L1 in foreign language classrooms. *Journal of Multilingual and Multicultural Development*, *41*(5), 406-419. https://doi.org/10.1080/01434632.2019.1684928

Sonbul, S., & El-Dakhs, D. (2020). Timed versus untimed recognition of L2 collocations: Does estimated proficiency modulate congruency effects? *Applied Psycholinguistics*, *41*(5), 1197-1222. https://doi.org/10.1017/s014271642000051x

Sun, H., & Fang, S. (2021). What constitutes vocabulary learning difficulty? A classroom-based study of semantic relatedness and L1 familiarity effects on L2 word learning. *Australian Journal of Applied Linguistics*, *4*(3), 82-102. https://doi.org/10.29140/ajal.v4n3.545

Tagashira, K., Kida, S., & Hoshino, Y. (2010). Hot or gelid? The influence of L1 translation familiarity on the interference effects in foreign language vocabulary learning. *System*, *38*(3), 412-421. https://doi.org/10.1016/j.system.2010.03.015

Terai, M., Yamashita, J., & Pasich, K. (2021). Effects of learning direction in retrieval practice on EFL vocabulary learning. *Studies in Second Language Acquisition*, *43*(5), 1116-1137. https://doi.org/ 10.1017/ S0272263121000346

Tsagari, D., & Giannikas, C. N. (2020). Re-evaluating the use of the L1 in the L2 classroom: Students vs. teachers. *Applied Linguistics Review*, *11*(1), 151-181. https://doi.org/10.1515/applirev-2017-0104

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352-373. https://doi.org/10.1037/h0020071

Tulving, E. (1974). Recall and recognition of semantically encoded words. *Journal of Experimental Psychology*, *102*(5), 778-787. https://doi.org/10.1037/h0036383

Tulving, E. (1976). Ecphoric processes in recall and recognition. In J. Brown (Ed.), *Recall and recognition* (pp. 37-73). Wiley.

Tversky, B. (1973). Encoding processes in recognition and recall. *Cognitive Psychology*, *5*(3), 275-287. https://doi.org/10.1016/0010-0285(73)90037-6

Urdaniz, R. P., & Skoufaki, S. (2019). Spanish L1 EFL learners' recognition knowledge of English academic vocabulary: The role of cognateness, word frequency and length. *Applied Linguistics Review*. https://doi.org/10.1515/applirev-2018-0109

Webb, S. (2007). Learning word pairs and glossed sentences: The effect of a single context on vocabulary knowledge. *Language Teaching Research*, *11*(1), 63-81. https://doi.org/10.1177/136216880607246 3

Webb, S., & Chang, A. C. S. (2015). How does prior knowledge affect vocabulary learning progress in an extensive reading program? *Studies in Second Language Acquisition*, *37*(4), 651-675. https://doi. org/10.1017/ s0272263114000606

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL - International Journal of Applied Linguistics*, *168*(1), 34-70. https://doi.org/10.1075/itl.168.1.02web

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing: A comparison of congruent and incongruent collocations. *Studies in Second Language Acquisition*, *35*(3), 451-482. https://doi.org/10.1017/s0272263113000107

Wolter, B., & Yamashita, J. (2015). Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics*, *36*(5), 1193-1221. https://doi.org/10.1017/s014271641400 0113

Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly*, *44*(4), 647-668. https://doi. org/10.5054/tq.2010.235998

Zheng, Y. (2011). Exploring Chinese EFL learners' vocabulary depth knowledge: The role of L1 influence. *The Journal of Asia TEFL*, *8*(3), 191-219.

(Received October 18, 2021; Revised March 18, 2022; Accepted March 18, 2022)

Appendix

Novel L2 Words

High					
Chinese	Familiarity rating	Real word	Pseudoword	No. of letters	No. of syllables
韭菜	6.79	lectern	lictern	7	2
山药	6.21	halcyon	holcyon	7	2
生姜	6.74	dulcet	dulket	6	2
扁豆	6.02	askew	aspew	5	2
苦瓜	6.24	puerile	puetile	7	2
莲藕	6.12	ursine	urkine	6	2
秋葵	5.62	yokel	yopel	5	2
香菇	6.64	turgid	tursid	6	2
木耳	6.52	credo	cremo	5	2
芦笋	5.24	elysian	elymian	7	3
Average	6.21			6.1	2.1
Low					
Chinese	Familiarity rating	Real word	Pseudoword	No. of letters	No. of syllables
扁琴	2.19	patois	patoin	6	2
黑管	2.00	nostrum	nostrim	7	2
铜钹	2.07	ribald	ribild	6	2
巴松	1.50	kismet	kasmet	6	2
板胡	1.90	simian	semian	6	2
三弦	1.90	vixen	vipen	5	2
排鼓	1.83	leonine	leonipe	7	2
响葫芦	1.67	zephyr	zophyr	6	2
短号	2.14	bovine	botine	6	2
沙槌	1.90	inured	inuted	6	2
Average	1.91			6.1	2.0
Mixed					
Chinese	Familiarity rating	Real word	Pseudoword	No. of letters	No. of syllables
麻雀	6.43	skulk	smulk	5	1
喜鹊	6.19	austere	auspere	7	2
海鸥	6.40	benison	beniton	7	3
燕子	6.50	laity	leity	5	2
乌鸦	6.43	quorum	quoram	6	2
山鹑	1.55	rotunda	rotonda	7	3
苍鹭	1.64	vassal	vassil	6	2
地鹃	1.79	mulct	mulat	5	1
松鸡	2.02	jocose	jacose	6	2
琴鸟	2.00	cadaver	cadamer	7	3
Average	6.39/1.80			6.1	2.1

