L1 and L2 lexical network of Chinese EFL learners: Focusing on the effects of semantic category relationship and L2 vocabulary knowledge

畢, 奇

(Degree)
博士（学術）

(Date of Degree)
2023-03-25

(Date of Publication)
2024-03-01

(Resource Type)
doctoral thesis

(Report Number)
甲第8533号

(URL)
https://hdl.handle.net/20.500.14094/0100482281

博士論文

L1 and L2 lexical network of Chinese EFL learners: Focusing on the effects of semantic category relationship and L2 vocabulary knowledge

（中国人英語学習者の L1 と L2 語彙ネットワーク形成における意味カテゴリー及び L2 語彙知識の影響）

2023 年 1 月

神戸大学大学院国際文化学研究科

BI QI （畢　奇）

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

1.1 Background of the Issue

It is said that vocabulary learning, especially learning the semantic information of vocabulary, is an area which has not yet received sufficient attention compared to grammar learning and other areas in the field of Second Language Acquisition (SLA). Imai (1993) criticized that foreign language education including English education seems to make light of vocabulary learning compared to grammar learning. Imai also criticized the general thought that the meaning of a word equals the explanation in the dictionary and vocabulary learning is just to remember the definition written in the dictionary.

In spite of vocabulary acquisition has received little attention and a general thought that the meaning of words equals explanations in the dictionary and vocabulary learning is just to remember the dictionary definitions (Imai, 1993). Some researchers tried to explain what should be learned when learning vocabulary. The vocabulary knowledge that needs to be learned is argued from many aspects, for example, phonetic or phonological information, lexical structure information, syntactic information, lexical collocation, network of associations, semantic information, semantic value, pragmatic information, vocabulary expansion with age, probability of encounter and so on (Akmajian, Demers, Farmer, & Harnish, 2010; Hiebert & Kamil, 2005; Loewen, 2015; Richards, 1976).

At the same time, some other researchers have also tried to figure out what the process is when learning vocabulary. In psycholinguistic area, language comprehension and production are said to depend on the activation of “mental lexicon” (lexical network structures in the brain) that contain linguistic information stored in long-term memory. Second language (L2) misuse at lexical and syntactic levels often occurs. The misuses of L2 learners are thought to be related not only to vocabulary size and depth of vocabulary knowledge, but also to the lexical network structure in the brain. However, it has not been clarified yet how the lexical network structure formed in the L2 learner’s brain, and what the relationship between the first language (L1) and L2 lexical networks in the brain is.

To know how the word relations and the differences in the range of semantic categories between L1 and L2 affect L2 vocabulary acquisition, Saji, Kajita, and Imai (2010) researched L2

Chinese learners' verb learning. Their results showed that there is a great difference between L2 Chinese learners and Chinese native speakers when using Chinese verbs to express same actions. An apparent influence of L1's semantic category or lexical knowledge to L2 verb usage was found in their research as the L2 Chinese learners' production patterns of Chinese verbs revealed to be in line with their own languages. In other words, the difference of semantic category between learners' L1 and L2 affect their L2 word learning to a large extent.

Jiang (2000) argued that among three interrelated aspects of the psycho-linguistic study of SLA: the study of representation, acquisition, and processing, representation has received little attention from researchers compared to acquisition and processing. For the study of L2 vocabulary acquisition, in particular, how L2 lexical information is represented in the mental lexicon has largely been ignored. Therefore, he outlined a psycho-linguistic model of vocabulary acquisition from the perspective of stages of lexical development in L2 acquisition known as three-stage L2 vocabulary acquisition model (see more details of this model in literature review **2.1**).

Focusing on the acquisition of "semantic information" in vocabulary acquisition, Bi (2020) investigated whether the semantic category relationship of Chinese and Japanese learners' L1 and L2 English influences their L2 vocabulary learning and found that in the acquisition of words with a broader semantic category in L1 than in L2, over-generalized misuse appeared, while in the acquisition of words with a broader semantic category in L2 than in L1, the broader parts of semantic information of L2 than L1 seemed easier to acquire for L2 learners. Therefore, the role of the semantic category relationship between L1 and L2 needs to be considered in L2 vocabulary acquisition. This study tried to further clarify the effects of semantic category relationship and the structure of lexical network of L2 learners.

1.2 Purpose of the Study

To know more details about the relationship between L1 and L2 words' semantic information and the structure of lexical network of L2 learners, and to verify the results in Bi (2020), the Experiment 1 in this study changed the perspective to try to investigate the lexical networks of some specific words whose relationships between the semantic categories of L1 and L2 are different, and to deduce the overall lexical networks in L2 learners' mental lexicon. Thus, Experiment 1 takes advantage of word association tasks, focusing on association responses' types

for each stimulus word and tried to clarify how the differences in semantic categories between L1 Chinese and L2 English are reflected in associations, to probe into the influence of L1 semantic categories on L2 vocabulary acquisition and to explore the mechanisms by which L2 misuse occurs for Chinese EFL learners.

L2 mental lexicon is considered not structured loosely or randomly but deeply related with both vocabulary size and the depth of vocabulary knowledge of each word. In word association tasks, the features of word knowledge of the stimulus words seem to influence the types of association words. Therefore, continuing the line of attempting to investigate the connections between lexical items and to know more details of the network structure of learners' mental lexicon, furthermore, to find some clues that provide information about the development process of acquiring vocabulary, the Experiment 2 in this study still takes advantage of word association tasks to study the effects of both breadth and depth of vocabulary knowledge on word association in L2 for Chinese EFL learners.

To gain insight into the way in which Chinese English learners represent their two languages in mental lexicon in more depth on the basis of the results of Experiment 1 and 2, an online collocation decision task was implemented in Experiment 3 to investigate whether or not and how the semantic category of L2 learners' L1 Chinese influence their L2 English vocabulary learning, and an experimental priming paradigm under masked condition were used to examine how semantically related and translation word pairs would be recognized. The accuracy and RT data collected in both tasks were analyzed. Then concerning the influence of semantic relationship observed in Experiment 1, a comparison of two conditions: semantic category relationship L1>L2 and the semantic category relationship L1<L2 are made to clarify the bilingual memory representation. Concerning the influence of vocabulary knowledge observed in Experiment 2, a comparison of HIGH and LOW proficient participants was made, too.

CHAPTER 2 PREVIOUS STUDIES

2.1 Bilingual Mental Lexicon Model and Lexical Networks in Mental Lexicon

2.1.1 An Overview of Bilingual Mental Lexicon Models

Language comprehension and production is considered to depend on the activation of "mental lexicon," which refers linguistic information (including morpheme, phoneme, syntax, semantics and so no) stored in one's long-term memory. L1 and L2 comparison lexicon research conducted from a cross-linguistic perspective is often considered as "bilingual lexicon" research. Word association and naming, recognition and recall, and language transfer and interference becomes the three major experimental approach dominated in this field. Much of the research concerning with the organization of a bilingual's two languages has been conducted up to now. The primary purpose has been to determine whether each language is represented separately (Kolers, 1963; Scarborough, Gerard, & Cortese, 1984) or both languages share a common memory store (de Groot & Nas, 1991; Kirsner, Smith, Lockhart, King, & Jain, 1984; Meyer & Ruddy, 1974). Further possibility is that there is a shared conceptual store between languages but in the same time, a separate lexical store is maintained for each language (Kroll & Stewart, 1994; Potter, So, Von Eckardt, & Feldman, 1984).

Some researchers in this area proposed some models to explain their research findings. According to French and Jacquet (2004), the representative four are the Word Association Model, the Concept Mediation Model, the Mixed Model, and the Revised Hierarchical Model. The Word Association Model argues that there are direct L1-L2 lexical links and L1-Concept links but not L2-Concept links. The Concept Mediation Model suggests that there is no L1-L2 lexical link but are L1-Concept links and L2-Concept links. French and Jacquet (2004) argues a shift from Word Association Model to Conceptual Mediation Model as second language proficiency increases. The Mixed Model insists that there are direct L1-L2 lexical links, L1-Concept links, and L2-Concept links, however, the lexical links do not disappear even proficiency increases. Furthermore, to solve the problems with task dependent differences in performance and different performance from L1 to L2 and from L2 to L1, the Revised Hierarchical Model (RHM) which weights bidirectional links were proposed.

2.1.2 The Revised Hierarchical Model (RHM)

Different from most models of monolingual language processing making a distinction of representations between form level (lexical) and semantic level, arguing strong mappings between L1 words' forms and their underlying meaning, the Revised Hierarchical Model proposed by Kroll and Stewart (1994), as an influential model of bilingualism, however, does not postulate such strong form-to-meaning mappings for L2 words. It is assumed that L2 words may only access semantics through their L1 equivalents.

The Revised Hierarchical Model describes the asymmetrical phenomenon that present in bilinguals' language representation in different language directions and tries to explain it. According to the RHM, the links between the L1 and concepts appear to be very strong, while the links between the L2 and concepts are weaker than that of the L1. Therefore, the conceptual access for L2 words might be mediated through L1 words, especially for low proficiency bilinguals. Although it was suggested that the strength of links between L2 and the concepts increase and this asymmetric lexical semantic organization disappear when bilinguals' L2 become more proficient.

Duyck and Brysbaert (2004, 2008) proposed an extension of the RHM and argues that translation is not a dual-route process that follows either the lexical or the semantic, but is the result of activation of both semantic and lexical representations. The strength of form-to-meaning mappings through which the activation is forwarded, is depending on both L2 proficiency (as in the RHM), and word types concerning not only the connection weights between the lexical and semantic level, but also the overlap between L1 and L2 lexical form representations.

2.1.3 The Bilingual Interactive Activation Model (BIA+) of Bilingual Memory Representation

The Bilingual Interactive Activation model (BIA+, Dijkstra & van Heuven, 2002) is a model of visual word recognition for bilinguals. The word frequency and familiarity are considered to affect and determine resting level activation levels of words (McClelland & Rumelhart, 1981). Since the familiarity of L2 is lower than that of L1, L2 with a low resting level activation takes longer to be activated. The difference from RHM is that it does not attribute the L2 lexical access delay to the indirectly pass by L1, but rather attribute to the L2 word has a lower resting-level

activation (Dijkstra & van Heuven, 2002; Gollan, Montoya, & Sandoval, 2008; Gollan, Slattery, Goldenberg, Van Assche, Duyck, & Rayner, 2011).

2.1.4 Jiang's Vocabulary Acquisition Model

Among the few researchers who have been concerned with semantic information learning in L2 vocabulary acquisition, Jiang (2000) as a representative, proposed a three-stage L2 vocabulary acquisition model. The model hypothesized that (a) firstly, L2 lemma information is copied from L1 into L2 lexical entry (word association stage), (b) then, other information like semantic, syntactic, and morphological specifications of L1 is transferred and integrated in L2 (L1 lemma mediation stage), and (c) lastly, the links between L2 words and their L1 translations become weak (full L2 integration stage). He also mentioned that the "full L2 integration stage" is just in principle while in practice the lexical development will stop at the second stage (L1 lemma mediation stage). This phenomenon is called "lexical fossilization" (pp. 50-55).

The claim in this model is in contrast to the assumptions held by many other L2 researchers who have insisted that L2 learners acquire new meanings while learning new L2 words. However, Jiang (2002, 2004) provided empirical support to his model.

Jiang (2000) stated that L2 words are linked to existing meanings transferred from L1 and that most words fossilize in the second stage (L1 lemma mediation stage) due to the constraints of L2 learners' learning situation: for example, sufficient input are difficult to get. Jiang also mentioned that even there are sufficient chances to expose the contextualized input, L2 acquisition may not always be reinforced. Since the LI lemma exists in the lexical entry, the contextualized exposure may also reinforce L1 lemma mediation by strengthening the connection between the L1 lemma information and the L2 lexeme information. Hence, it is almost impossible for L2 learners to achieve the last stage (full L2 integration stage).

Moreover, Jiang (2004) mentioned that one of the factors needs to be considered is the learners' proficiency. Because the participants in Jiang's study were students who had lived in the United States for a limited number of years, it is not clear yet whether the "lexical fossilization" would still happen or disappear in the case that L2 learners who have lived in English environment much longer and who are much more proficient.

2.2 Semantic Category Relationship between L1 and L2

In Jiang's three-stage model, all L2 words were considered to have the L1 lemma mediation stage when L1's lemma information is transferred into L2 lexical entry and affects L2 words use, whether their semantic category is larger or smaller than L1. The relationship between L1 and L2 in terms of semantic category was not taken into consideration. In the research field of SLA, referring to White (1989)'s theory, the different relationships of grammatical rules between L1 and L2 lead to different situation of L2 grammatical acquisition. It is considered that when the rule of L1 is less general than L2, only positive evidence is enough to bring about exact L2 acquisition (an expansion of meaning in L1).

White (1989) used one of the examples of the acquisition of the complements of verbs to explain this problem. The verb "see" can take a direct object noun phrase or a sentential complement, as follows: (a) Marry saw John, and (b) Marry saw that John was tired. When a L2 learner who has already known that "see" takes a direct object noun phrase but does not know that it can also take a sentential complement notice sentences like (b) in the input process, this positive evidence can bring about the change of recognition of the rules of the verb "see" for L2 learners (White, 1989, p.142). Conversely, when L1 contains items or rules which are incorrect for L2 and which must be eliminated, explicit negative evidence may be required otherwise it will result in over-generalization. To explain more clearly, we still cite one of the examples in White (1989). Although adverb placement is free in English, adverbs cannot intervene between a verb and its direct object. For example, sentences like "Mary slowly does her homework.", "Mary does her homework slowly.", "Slowly Mary does her homework.", and "Mary is slowly doing her homework." are all grammatical sentences in English. "Mary does slowly her homework. *" is not grammatical in English, however. In French, the equivalent of "Mary does slowly her homework. *" (Marie fait lentement ses devoirs.) is also grammatical. Therefore, when a French learns about this grammar rule of English, explicit negative evidence will be required to draw to the learner's attention to the fact that sentences like "Mary does slowly her homework. *" are not permitted in English. (White, 1989, p.142., "*" means ungrammatical here.)

Considering that the idea of White (1989) may be also appropriate for vocabulary acquisition, Bi (2020) focusing on the acquisition of "semantic information" in vocabulary acquisition, argued that White's theory may be also appropriate for vocabulary acquisition and that the semantic

category relationships between L1 and L2 need to be considered. Bi (2020) assumed that when L2 semantic category is smaller than L1, over-generalization will appear, and the L1 lemma mediation stage will continue until the negative evidence is presented and acquired. On the other hand, when L2 semantic category is larger than L1, only positive evidence is enough, so the L1 lemma mediation stage may not be existent or really short, and the L2 integrated stage when L2's lemma and lexeme information are integrated in the lexical entry is easy to enter. Bi (2020) investigates whether the semantic category relationship of Chinese and Japanese learners' L1 and L2 English influences their L2 vocabulary learning by implementing a writing test and a judgment test and analyzing difference between learners with different English proficiency. As a result, the effect of the semantic category relationship between L1 and L2 in L2 vocabulary acquisition were confirmed. For words with a broader semantic category in L1 than in L2, over-generalized misuse appeared, while for words with a broader semantic category in L2 than in L1, the broader parts of semantic information of L2 than L1 seemed easier to be acquired for L2 learners.

There are some other studies provided empirical support for the presence of influence by the semantic category of L2 learners' L1 to their L2 vocabulary learning. Jiang (2002, 2004) researched vocabulary semantic categories with Chinese and Korean learners of English, while Saji, Kajita, and Imai (2010) studied vocabulary semantic categories with Japanese and Korean learners of Chinese.

2.3 Research of Word Association

According to Fitzpatrick and Thwaites (2020), research on word associations has developed a lot since its appearance in the late nineteenth century. They paid attention to the word association in applied linguistics context, summarized its development path from psychoanalytic research in the early twentieth century, to broader psychological work including child cognition around the 1960s, to applied linguistics research including L2 investigations more recently. The use of word association was concluded as to investigate "knowledge of vocabulary items" and "word storage and retrieval."

The research of L2 lexical knowledge has been an attractive area for researchers who are interested in how L2 learners know about words and what the structures are the lexical networks in their mental lexicon. According to Jenkins (1954), "a first approach to the examination of

interrelationships between word units may be found in the classic word association test." (pp. 112-118). In this context, attention mainly focusing on using the word association tasks as a method to know the contents and structures of the L2 lexicon by investigating what kind of L2 word knowledge participants have. Read's (1993) Word Associates Format in which test takers have to find relations between a stimulus word and a given set of other words is based on the idea of vocabulary networks. Nation's word knowledge framework (Nation, 1990; 2001) and Richards' lexical competence assumption (1976, pp.79-81) use the word "association" and define its notion explicitly: Nation lists "associations" as one of nine components of "knowing a word," and Richards put "knowledge of the network of associations between a word and other words" as one of eight characters of lexical competence. Both of them focused mostly on semantic associations. Quality of vocabulary knowledge was considered as a feature of the semantic network by some researchers (Aitchison, 2012; Henriksen, 1999). They suggest that learners acquire the meaning of words by using their relations to other words, acquiring more related words leads to deepening of word knowledge, and the lexical networks become denser and highly interconnected. Therefore, among the wide use of word association research, the word association task sometimes behaves as a common vocabulary test and is used as an approach to assess vocabulary knowledge. However, Cremer, M., Dingshoff, de Beer, and Schoonen (2011) drew a conclusion that word association tests are unsuitable as assessment tools of word knowledge according to the result of their experiment which made a comparison of word association responses produced by native speakers and L2 learners. Some other researchers also suggest that word association relates to not only word knowledge but also other aspects of language proficiency. Moreover, some researchers argued that similarity of word association responses to L1 could distinguish between native speakers and L2 learners, but could not distinguish L2 learners with different proficiency level (Zareva, Schwanenflugel, & Izura, 2005; Zareva & Wolter, 2012).

As a way to know the structure and contents of L2 lexicon, the following points are where the attention of word association researchers have been mainly paid to: the influence of stimulus words, including the influence of stimulus words' features, like frequency (for example, highly frequent words lead to native like responses, in Fitzpatrick, 2007), familiarity, if a stimulus word is abstract or specific (concrete words elicit more homogeneous responses than abstract words in De Groot,1989), the word class (especially word's part of speech) (Cronin, 2002; Nissen &

Henriksen, 2006; Higginbotham, 2010. see more detail in **2.3.2**) and so on; the influence of participants (the influence of participants' L2 proficiency, age and so on); the categorization of association responses (see more detail in **2.3.1**); the L2 response similarity to their L1 response (a comparison of L2 learners L1 and L2, see more detail in **2.3.3**); and the L2 learners' response similarity to native speakers' (a comparison of L2 learners and L1 speakers, in other words, native-like degree). The number of responses produce by language learners are commonly used as an index in these comparisons. Besides focusing on the number of association responses, multiple methods including individual and continuous association tests, response time measurement (De Groot, 1989; Izura & Playfoot, 2012; Van Hell & De Groot, 1998 a, b), association strength measurement (Nelson, McEvoy, & Dennis, 2000) and analysis of response homogeneity were also considered. Some researchers used some or all these methods (Fitzpatrick & Izura, 2011; Playfoot & Izura, 2013; Santos, Chaigneau, Simmons, & Barsalou, 2011), but little progress has been made in a better understanding of psycho linguistic implications.

About the influence of stimulus words' frequency, Fitzpatrick (2006) suggested that stimulus words which elicit very homogeneous sets of responses are relatively not ideal to researchers because the differences are difficult to emerge in this situation, which explained by Fitzpatrick (2006) as such stimulus words "mask more subtle behavior patterns" (Fitzpatrick, 2006, p.127). Fitzpatrick improved this problem by selecting cues for several of her studies (Fitzpatrick, 2006, 2007, 2009) from a specific academic word list to exclude some most frequent words in English, which as a result allows researchers to avoid the problem of using too many high frequency stimulus words.

Among these points, about the influence of participants' L2 proficiency, some research found that when the L2 proficiency becomes high, there is an increase in the number of responses produced by stimulus words in a multiple-response word association task (Lambert, 1956; Zareva, Schwanenflugel, & Nikolova, 2005), responses that are the same as those of L1 speakers (Schmitt, 1998a; Fitzpatrick, 2012), and paradigmatic responses or meaning-based responses (Orita, 2002; Zareva, 2007; Zareva & Wolter, 2012; Khazaeenezhad & Alibabaee, 2013).

Some representative research regarding the categorization of association responses, the influence of stimulus words' part of speech, and the L2 response similarity to their L1 response (a comparison of L2 learners L1 and L2) are introduced in more detail.

2.3.1 The Categorization of Association Responses

Categorization of associations responses were seriously considered. Cremer, Dingshoff, de Beer, and Schoonen (2011) referred to De Groot (1980) which focused on meaning-based responses and distinguishes objective (superordinate, subordinates, and synonyms) versus subjective (e.g., vase-mother) semantic relations. In addition, the almost most popular way of categorization was the syntagmatic-paradigmatic system, which can date back to Saussure (earliest in 1916, cited by Fitzpatrick & Thwaites, 2020 referring Saussure, in Bally & Sechehaye, 1966), and used and developed by many researchers (e.g. Batty, 2012; Cremer, Dingshoff, de Beer, & Schoonen, 2011; Cronin, 2002; Dronjic & Helms-Park, 2014; Ervin, 1961; Entwisle, Forsyth, & Muuss, 1964; Håkansson & Norrby, 2010; Li, 2011; Jiang & Zhang, 2019; Khazaeenezhad & Alibabaee, 2013; Namei, 2004; Nissen & Henriksen, 2006; Norrby & Håkansson, 2007; Orita, 2002; Read, 1993, 2000; Wolter, 2001; Yokokawa, Yabuuchi, Kadota, Nakanishi & Noro, 2002; Yabuuchi (2015); Zareva, 2007; Zareva & Wolter, 2012; Zhang & Koda, 2017). Paradigmatic responses refer to responses in the same grammatical class as stimulus words which could replace the stimulus words in a specific sentence, while syntagmatic responses refer to responses in a different grammatical class as stimulus words which are always in the position before or after the stimulus words in a specific sentence.

However, numerous problems in this way of categorization identified and the way of categorization has received criticism since the 1970s (Deese, 1966 referred by Fitzpatrick & Thwaites, 2020; Fitzpatrick, 2006, 2007; Higginbotham, 2010; Meara, 2009: 22; Nissen & Henriksen, 2006; Petrey, 1977). According to Fitzpatrick and Thwaites (2020), Deese (1966) gives “grass-green” and “sky-blue” as examples of “syntagmatic responses which are not so much representative of the sequential transitions in ordinary language, as they are of the most basic and important cognitive characteristics of grass and sky”. Petrey (1977) also gives “add-flour” (syntagmatic), and “add-cook” (paradigmatic) as examples, and explained that both “flour” and “cook” come from the same situation and are related in the similar way to “add”, but would be classified as opposites in this way of distinction. Thus “syntagmatic responses properly so-called are highly problematic because either semantic or episodic storage could produce them” (Petrey, 1977, p.66). Nissen and Henriksen (2006) criticize the syntagmatic-paradigmatic shift as they

found there was a surprisingly strong native speaker preference for syntagmatic responses. Fitzpatrick (2007) suggests that the syntagmatic-paradigmatic distinction may not be exact enough to provide informative data about individual participants and their responses. Fitzpatrick's own work with a more detailed scheme supports her claims. Higginbotham (2010, p.383), contrasting the syntagmatic-paradigmatic scheme with that of Fitzpatrick, suggests that the use of Fitzpatrick's way of categorization can make some obscured details in syntagmatic-paradigmatic distinction clear. Higginbotham (2010) found that in addition to supporting previous findings regarding differences between L1 and L2 associations, such as the greater heterogeneity of L2 responses, Fitzpatrick's new way of categorization could reveal significant differences in the elicited responses types. For example, L2 participants gave more conceptual associations, while L1 participants were more likely to elicit "defining synonyms." As Fitzpatrick pointed out (2006), these responses would all have been coded as paradigmatic under the syntagmatic-paradigmatic classification, thus obscuring this kind of difference. Moreover, she called for using both broad and specific subcategories to make sure more responses are categorized reasonably.

Fitzpatrick (2006) tried something new and used four main categories: meaning-based, position-based, form-based, and erratic associations. Each main category was divided into detailed subcategories to classify the relation between stimulus words and association responses, as only general main categories do not allow a very precise method of categorization.

Fitzpatrick and Izura (2011) extended Fitzpatrick (2006)'s categorization and created two additional categories (form and meaning, e.g., pencil-pen; and meaning and collocation, e.g., pen-paper) for responses associated to the stimulus word by more than one link. This categorization system includes the following six categories: form and meaning, meaning and collocation, collocation, form, equivalent meaning, and nonequivalent meaning. The coding scheme used to categorize the responses in present study is adapted from the categorization systems used by Fitzpatrick (2006, 2007, 2009) and Fitzpatrick and Izura (2011).

2.3.2 The Influence of Stimulus Words' Part of Speech

The feature of stimulus words used influence the association responses elicited. Many researchers found the influence of word's part of speech on response types (Cronin, 2002; Higginbotham, 2010; Komori, 2015; Nissen & Henriksen, 2006). Cronin (2002) and Nissen and

Henriksen (2006) found that nouns, for example, elicit paradigmatic responses more often than verbs or adjectives do. Nissen and Henriksen (2006) found that there was no significant difference between syntagmatic responses given to nouns in the L1 and L2. In other words, the results showed a surprisingly majority of syntagmatic responses also in L1. Nissen and Henriksen (2006)'s discussion of this finding raises many questions about the basic nature of word associations which have rarely be discussed before. They suggested that "words do not possess equal potential for yielding similar proportions of syntagmatic and paradigmatic responses in word association tests" (Nissen & Henriksen, 2006, p.403). Acquisition difficulty according to the more abstract semantic features of verbs and adjectives than nouns, the low possibility of isolated appearance and the custom of learning verbs with nouns as a set were mentioned as several aspects of the cognitive nature of verbs and adjectives which easily lead to syntagmatic responses. Furthermore, Nissen & Henriksen (2006) implied that roles of different part of speech words in sentences may influence word association responses. This is a vital point, not only because it provides explanation for stimulus of different part of speech elicit different proportions of syntagmatic and paradigmatic responses, but also because it put forward the semantic basis of word association by suggesting that co-occurrence might be an influential factor of association.

Komori (2015)'s results suggested that nouns elicit more meaning-based association in both L1 (Japanese and English) and L2 (English). Higginbotham (2010) also reported that nouns elicit more meaning-related responses than stimulus words in other part of speech.

Therefore, when comparing the results of different studies, the inherent difference between part of speech is one aspect needs to be taken into consideration. In the present study, the list of stimulus words consisted of both nouns and verbs.

2.3.3 The L2 Response Similarity to L1: A Comparison of L1 and L2 Mental Lexicon

Some representative research regarding the L2 response similarity to their L1 response (a comparison of L2 learners' L1 and L2) are introduced.

Yokokawa, Yabuuchi, Kadota, Nakanishi, and Noro (2002) examined Japanese native speakers' semantic organization of the lexicon in both L2 English and L1 Japanese by using the meaning-based word association task for the meaning familiar stimulus words with fine-grained semantic categories. The results showed that there were more syntagmatic responses than

paradigmatic responses for abstract-concrete stimulus words, however, more paradigmatic responses than syntagmatic responses for categorical stimulus words. They also found that there was a high degree of homogeneity between L1 and L2 association patterns but the degree of mating between L1 and L2 association responses was about fifty to sixty percent and the qualitative analysis was made to see the differences in detail in their study.

Yabuuchi (2015) gave Japanese and English stimulus words to Japanese native speakers to compare the Japanese speakers' L1 and L2 responses. She focused on the number of responses, and found that there was no difference on superordinate relationship responses (e.g., animal-dog, cat, elephant, et al.), but a difference in responses for abstract and concrete stimulus words. The association responses for concrete stimulus words are similar in L1 and L2, however, for abstract stimulus words the results are diversified.

Nissen and Henriksen (2006) gave English and Danish stimulus words to native speakers of Danish and focused on word's part of speech as the prime variable, they suggested that factors such as frequency, concreteness, and mental image may have also influenced the word association results and tried to limit the influence of these potential variables. The results showed that the part of speech of the stimulus words has influenced the participants' word association responses, both in L1 and L2. Nouns elicited more paradigmatic responses than verbs and adjectives in both L1 and L2. However, no significant difference was found between syntagmatic responses given to nouns in the L1 and L2. In other words, the results showed a surprisingly majority of syntagmatic responses also in L1.

Komori (2015) gave Japanese and English stimulus words to Japanese native speakers and gave same English stimulus words to English speakers to compare the results of association responses. She used Fitzpatrick (2006)'s way of categorization and found that for Japanese native speakers, there were more position-based association and less form-based association in Japanese than in English. This is consistent with the result of Nissen and Henriksen (2006).

Fitzpatrick and Izura (2011) gave Spanish and English stimulus words to Spanish native speakers and compared word association responses in L1 Spanish and L2 English from the perspective of responses types and response times. The result showed that generally, L2 responses are fewer and slower than L1 responses. The category RTs is broadly similar between L1 and L2, and participants took more time but most frequently to produce nonequivalent meaning

associations in both languages. They also found the difference between L1 and L2 response times became smaller as proficiency increased.

2.4 Methodological Issues of Word Association Task

The number of responses requested for each stimulus word is one of the notable methodological issues. In researches up to now, the number of responses requested for each stimulus word has been varied. Some researchers prefer the single-response procedure as the first responses are considered have the strongest connection with stimulus words (Nelson, McEvoy, & Dennis, 2000). Nelson, McEvoy, and Dennis (2000) found that in the second responses, words which have weak connections with the stimulus words appear. Moreover, they found that when the primary association was not produced as the first response, it tended not to be produced on the second. Furthermore, the first response may act as a new stimulus for the later responses thus eliciting a chain of responses, this response chain sometimes elicits some responses which 'are not independent of the first' (Nelson, McEvoy, & Dennis, 2000, p.891) and is considered to negatively affects the reliability of association responses. Playfoot, Balint, Pandya, Parkes, Peters, and Richards (2016)'s study also suggested that the primary association responses generated are actually the first word that comes to participants' mind in the free word association task. On the other hand, other researchers advocate eliciting multiple responses as it allows a greater degree of access into the lexicon of learners (De Deyne, Navarro, & Storms, 2012; Schmitt, 1998b). Schmitt (1998b) used word association responses to determine whether a participant's associative network is native-like. Schmitt questions whether participants always produce responses from the strongest connection ones to least in sequence (1998, p.391). He argues for multiple responses to one stimulus word as in this way we can 'better capture the richness of a subject's associative network,' which can then be used to have a comparison to a native speaker's network. As one of the purposes of this study is to compare the network of participants' L1 and L2, the multiple response paradigm is adopted.

Moreover, word association responses tend to vary depending on the way stimulus words presented to participants and task demands (in oral or written form), too. For example, Palermo (1971) compared responses using the aural-oral and written-written methods for native speaker children and found the mean proportion of paradigmatic responses was higher when the test was

administered using the aural-oral method. Suzuki-Parker and Higginbotham (2019) compared responses using the aural-oral and written-written methods for Japanese learners of English and found significant differences between the two methods. The biggest difference was found between the responses in the meaning-based category: the meaning-based responses were more produced when the test was administered by a written-written method than by an aural-oral method. The present study adopted a written way of stimulus words presentation and association responses collection for the focus on semantic category and the practical reasons as the convenience of implementation.

In addition to the categorization of association responses as mentioned above, the evaluation of responses is another issue. Whether the rater is L1 or L2 speakers and whether is the participant who elicits the responses may lead to different results. Fitzpatrick (2006) concerned the issue that to what extent the researcher's understanding are consistent with the respondents' ideas. She noted that even use a newly improved method to categorize responses, there would still be some items whose connection between stimulus words and responses is difficult to consider. The solution here was to conduct a retrospective interview to respondents in order to know their ideas better. This solution is somewhat problematic as there is a possibility that the association procedure is automatic and unconscious. Thus, interviews conducted later would reveal little of actual thinking processes underlying association response. This therefore raises challenging questions about how to determining the exact responses producing process.

2.5 Vocabulary Knowledge and Word Association

2.5.1 Vocabulary Knowledge

When talking about learning vocabulary, the first thing to do first is to define what the vocabulary knowledge is. As for what knowledge of a word consists in, Cronbach (1942) summarized in items of generalization, application, the breadth of meaning, precision, and availability in the study of diagnostic vocabulary test of native language and it brought definition of vocabulary knowledge into focus as well. Richards (1976) proposed eight assumptions concerning the nature of lexical competence further, as including vocabulary expansion with age, probability of encounter, function and situation, syntactic behavior, underlying form and the derivations, network of associations, semantic value and different meanings which included most

aspects of vocabulary knowledge. However, as Meara (1996) commented that "most of the current research still looks at vocabulary as discrete items", more sensitive and convincing models concerning learning's personal experience need to be developed. Hiebert and Kamil (2005, p.70) discussed the complexity of word knowledge referring to Nation (1990) which identified that there are eight different aspects of word knowledge, including knowledge of a word's spoken form, its written form, the way it acts in sentences, lexical collocation, its frequency in both oral and written language, its conceptual meaning, its register (about when and how it is commonly used), and its association with other words. Akmajian, Demers, Farmer, and Harnish (2010, pp.14-15) argued that information of words is complex and it could be summarized as phonetic/phonological information, lexical structure information, syntactic information, semantic information, and pragmatic information. Loewen (2015, p. 98) mentioned the concept of depth of knowledge and breadth of knowledge. He indicated that in many aspects of vocabulary knowledge, depth of knowledge includes grammatical function, collocation, frequency, register, domain and so on, while breadth of knowledge means the number of words that learners know. Nation (2001) pointed out that item knowledge and system knowledge are different and stated that learning vocabulary itself is not the only aim, vocabulary learning could help learners listen, speak, read, and write effectively. He also insisted that vocabulary test needs to differentiate vocabulary knowledge and vocabulary use.

2.5.2 Influence of Vocabulary Knowledge to Word Association Patterns

There was much debate on the issue of influence of vocabulary knowledge to word association patterns. Some studies have demonstrated that L2 learners experience a syntagmatic-paradigmatic shift. This means that in the initial stages of the learning process, learners usually produce more form-based and more syntagmatic associations; then along with the development of language proficiency, they tend to produce less form-based and more paradigmatic associations (e.g., Söderman, 1993 referred in Fitzpatrick, & Izura, 2011; Sökmen, 1993). However, other studies have argued that this well-known syntagmatic-paradigmatic shift may be affected by participants' vocabulary knowledge of stimulus words (Fitzpatrick, 2006, 2007; Namei, 2004; Nissen & Henriksen, 2006; Wolter, 2001). For the fact that when the familiarity of a stimulus word is low, even native adult speakers whose language proficiency are usually high tend

to produce more form-based associations, Wolter (2001) put a particular emphasis on that the vocabulary knowledge level of stimulus words, instead of the general language proficiency leads to the variation of word association patterns. Namei (2004) also found that both the L1 and L2 mental lexicons are organized along a kind of word-knowledge continuum and the syntagmatic-paradigmatic shift is different for each individual word. Words may linked with others in form in the initial unfamiliar stages, then, they linked syntagmatically with some other words, and linked paradigmatically with more words lastly as become well-known. In a word, the development of vocabulary knowledge influence the word links in mental lexicon which reflected in the variation of word association patterns.

As for the research of the relationship between vocabulary size and the word association, in Kikuchi, Yamamoto, Yoshimura, Yabuuchi, and Tanimura (2001), the L1 Japanese and L2 English word association tests and a standardized test of L2 vocabulary size, Nation (1990)'s Vocabulary Levels' Test (VLT) were implemented to Japanese undergraduate students to probe how are the association tests both in L1 and L2 correlated with VLT. The results found that the correlation coefficients were statistically significant but relatively low between the two association tests and Nation's VLT. Further researches are expected to know the influence of vocabulary knowledge to word association patterns.

2.6 Using Priming to Measure Lexical Organization and Some Methodological Problems

The semantic priming is known as one of the most commonly used way to explore the structure of the mental lexicon in cognitive psychology research field. It is thought that when we see a word, the word's meaning along with words that are semantically related or associated are activated (Collins & Loftus, 1975; Collins & Quillian, 1969). The semantic priming effect has been observed in many previous experiment (see McNamara & Holbrook, 2003; Neely, 1991 for reviews). To know how words are represented in memory more accurately, several experimental constraints are necessary.

Frequency, length of prime and target words, and use of strategies (e.g., an expectancy strategy) must be taken into consideration (Balota & Chumbley, 1984, 1985; Chumbley & Balota, 1984; Neely, Keefe & Ross (1989). Methodological factors, as the Relatedness Proportion (RP, the proportion of related prime-target trials out of all prime-target word trials), the Stimulus Onset

Asynchrony (SOA, the time delay between the onset of the prime and the onset of the target), and the NonWord Ratio (NWR) influence whether the strategy is used by participants (Neely, 1991, Altarriba & Basnight-Brown, 2007). It has been shown that the semantic priming effect is larger when the RP is larger (i.e., more related word pairs than unrelated overall in an experiment, de Groot, 1984). Therefore, it is better to keep RP low when designing priming experiments. It has been shown that if the SOA is too long, participants will have a long time to think about the prime word and may lead to a quick response. SOA lengths of less than 300 msec at least or even shorter are considered to be ideal (Hutchison, Neely, and Johnson, 2001). The NWR equal .50 is ideal (Neely, Keefe, & Ross, 1989) for participants. The word response may be biased to choose Word when the NWR is below .50, but biased to choose Nonword when the NWR is over 0.50 (McNamara & Holbrook, 2003). Therefore, it is important to take these issues into account to dealing with the use of strategic processes.

2.7 Translation and Semantic Priming in Bilinguals

Many previous studies conducted on bilingual language storage and representation in memory. Some have used the cross-language semantic priming paradigms (e.g., Chen & Ng, 1989; Frenck & Pynte, 1987), and others used the translation priming paradigms (Altarriba, 1992; Williams, 1994).

The data reported in these studies seems discrepant on account of methodological differences (a review, see Altarriba & Basnight-Brown, 2007). For example, strong semantic priming effect has been observed in both L1-L2 and L2-L1 directions for some cross-language priming experiments (Basnight-Brow, & Altarriba, 2007; Keatley & de Gelder, 1992; Schoonbaert, Duyck, Brysbaert, & Hartsuiker, 2009), whereas some other research has observed the priming effect only in the L1-L2 direction (e.g., Williams, 1994), or even failed to gain any significant semantic priming results (Grainger & Beauvillain, 1988).

Translation priming studies also have revealed a variation of results. Strong translation priming effects in both the L1-L2 and the L2-L1 directions have been observed in some studies (concept featural model (DCFM) in De Groot, 1992; De Groot, & Poot, 1997; Basnight-Brown & Altarriba, 2007; Kiran & Lebel, 2007) as translation priming symmetry; but not in either direction or only in one in other studies (in L1-L2 direction only in Altarriba, 1992; BasnightBrown &

Altarriba, 2007; Chen, Zhou, Gao & Dunlap, 2014; Finkbeiner, Forster, Nicol, & Nakamura, 2004; Gollan, Forster, & Frost, 1997; Grainger & Frenck-Mestre, 1998; Jiang, 1999; Jiang & Forster, 2001; Schoonbaert, Duyck, Brysbaert, & Hartsuiker, 2009; Wen & van Heuven, 2016; some explorations in L2-L1 direction, in Finkbeiner, Forster, Nicol, & Nakamura, 2004; Jiang & Forster, 2001; Nakayama, Ida, & Lupker, 2016)) as translation priming asymmetry.

It seems that information of participants in these previous literatures were not enough, so the discrete results of priming effects reported until now in the cross-language semantic and translation priming may due to the different types of bilinguals have participated (Altarriba & Basnight-Brown, 2007). In addition, even when types of bilinguals are not that different, bilinguals' language experiences may affect, too. Further researches are expected to clarify these problems.

CHAPTER 3 THE PRESENT STUDY: EXPERIMENT 1

Word association in L1 and L2 for Chinese EFL learners: Focusing on response types, and influence of L1 semantic category on L2 vocabulary acquisition

3.1 Purpose of the Experiment 1

The present experiment attempted to investigate the features of lexical networks in both L1 and L2 of Chinese learners of English who learn English as a foreign language (EFL) by analyzing the number and types of word association responses in word association tasks. This experiment tried to investigate how it is different or similar in association responses between L1 and L2, between the association responses produced earlier and later, and between stimulus words with different part of speech. Furthermore, the present experiment focused on whether the results of word association differ according to the relationship of the semantic categories between L1 and L2. In other words, is there some difference between the results of words like “drink” which in L1 Chinese have a broader semantic category and the results of words like “wear” and “break” which in L2 English have a broader semantic category (See Bi, 2020).

3.2 Research Questions

The research questions of present study were posited as follows:

Research Question 1: Is there a significant difference of the number of association responses between L1 and L2? Are some response types more numerous than others?

This question will be addressed in both the L1 (Chinese) and the L2 (English).

Research Question 2: Is there a decrease of the number of association responses from the first association to the twentieth in both L1 and L2?

Research Question 3: Does the stimulus words, especially their part of speech (verbs or nouns) affect the response types?

The last but not the least, as for the influence of semantic category relationship of L1 and L2 to L2 word (collocation) acquisition, the research question was set as follows.

Research Question 4: Does the result of word association differ according to the relationship of the semantic categories between L1 and L2? In other words, is there difference between the results of words like “drink” which in L1 (Chinese) have a broader semantic category and the

results of words like “wear” and “break” which in L2 (English) have a broader semantic category?

3.3 Hypotheses

The research questions of present study were posited as follows:

Hypothesis 1: Although it is reasonable to assume that the number of L1 responses is larger than that of L2 responses, as the fact that the vocabulary size of L1 is apparently larger than L2 for EFL learners. According to some previous research like Nissen and Henriksen (2006), and Wolter (2001), the association situation is always similar in L1 and L2. If L2 lexical network formed on the basis of L1, the number of association responses in L1 and L2 may be similar. Even some response types will be more numerous than others, the number of same association types may be similar in L1 and L2 if the L2 lexical network formation is affected by the L1.

Moreover, according to Fitzpatrick and Izura (2011), there are several different kinds of potential activation triggers (meaning, collocation, form) and that the activation event is stronger when more than one of these triggers are present. Therefore, when there are dual-links between stimulus and responses (meaning and collocation, for example) the response may be activated easier when there is only a single-link (collocation only, for example).

Hypothesis 2: Referring to Yokokawa, Yabuuchi, Kadota, Nakanishi, and Noro (2002) which researched the L1 and L2 word association of Japanese EFL learners, there will be a decrease of the number of association responses from the first association to the twentieth in both L2 and L1. The types of response are assumed to be different in former elicited words and later ones. This means that the strength of links between words are different by types. If the patterns are similar in L1 and L2, there may be a similar lexical network between L1 and L2. However, if the patterns are different, lexical network between L1 and L2 may be qualitatively different.

Hypothesis 3: Many researchers have found an influence of word’s part of speech on response types mentioned in the previous studies (Cronin, 2002; Higginbotham, 2010; Komori, 2015; Nissen & Henriksen, 2006). For example, Cronin (2002) and Nissen and Henriksen (2006) argued that nouns elicit paradigmatic responses more often than verbs or adjectives do. Therefore, we assume that the stimulus words’ part of speech (verbs or nouns) will affect the response types and nouns will elicit paradigmatic responses more than verbs do. However, if there is a shift from syntagmatic network formation to paradigmatic network formation as the vocabulary size

increases in L2, similar as in L1, more paradigmatic responses in L1 may be observed than in L2. However, for the environment of EFL, this shift may cannot be observed in L2.

Hypothesis 4: The results of word association will be different according to the relationship of the semantic categories between L1 and L2. According to Bi (2020), words like “drink” which in L1 (Chinese or Japanese) have a broader semantic category are considered to be more difficult to acquire than words like “wear” and “break” which in L2 (English) have a broader semantic category. Thus, the association verb “drink” for stimulus words “medicine”, “soup” and “porridge” may appear in L2, and may even be more than “take” or “eat”. On the other hand, for the stimulus word “drink”, nouns like “medicine”, “soup” or “porridge” which composes improper collocations with “drink” in L2 may appear as the association responses. However, for the stimulus word “wear”, besides “coat”, “jeans” and “shoes,” which compose collocations with “wear” in L2 just same as “大衣”, “牛仔裤” and “鞋” which compose collocations with “穿” in L1, there may be nouns like “hat”, “necktie”, “watch”, “glasses” and so on which can compose collocations with “wear” in L2 but the equivalents in L1 are not acceptable as collocations with “穿”. Similarly, for the stimulus word “break”, association responses of both stimulus words “打破” and “弄坏” are assumed to be associated.

3.4 The Main Study

3.4.1 Participants

Thirty-seven Chinese university students (twenty females, sixteen males and one without information) majoring in stomatology at first or third year took part in the study. However, two did not complete the tasks as required and produced L2 responses in L1 association tasks, were excluded from the study. Thus, the valid data become thirty-five (nineteen females, fifteen males and one without information). The thirty-five participants learn English as a foreign language for 9.4 years in average. Participants were an average of 18.6 years old (range: 17—23). None of them have an experience of going abroad except one who had been to the United Kingdom for two months.

3.4.2 Materials

Considering the semantic category relationship between L1 Chinese and L2 English, the

word “drink” was considered as representative of English semantic category narrower than Chinese “喝 he”. The two words “wear” and “break” were considered as representative of English semantic category broader than Chinese. In L2 association task, verbs “drink”, “wear” and “break”, nouns “medicine”, “soup” and “porridge” that compose collocation with “drink” which showed significant difference between high and low proficiency learners in Bi (2020), and nouns “coat”, “jeans” and “shoes” that compose representative collocation with “wear” were used as stimulus words. The reason to use not only “drink”, “wear”, “break”, but also words that compose collocation with them as stimulus is that in this way, we can better investigate the links between them in mental lexicon in both directions. In L1 association task, the equivalents of L2 stimulus words were used. All words and the equivalent relationship are showed in Table 1.

Table 1

Stimulus Words in L1 and L2

L1 (Chinese)	喝	穿	打破、弄坏	药	汤	粥	大衣	牛仔裤	鞋
L2 (English)	drink	wear	break	medicine	soup	porridge	coat	jeans	shoes

Table 2 showed that “drink” in English has a narrower semantic category than “喝” in Chinese. For “soup”, “medicine”, and “porridge”, it is possible to say“喝汤”“喝药”“喝粥” in Chinese, but it is not acceptable to say “drink soup”, “drink medicine”, “drink porridge” in English. The collocation used in this study used bold font (Table 3 and Table 4 in the same way).

Table 2

Collocation Information of "drink（喝）" in L1 and L2

	英語	中国語
water 水 wine 葡萄酒 beer 啤酒 milk 牛奶	**drink**	**喝**
soup 汤	eat	
medicine 药	take	**喝**（吃）
porridge 粥	have, eat	**喝**（吃）

Table 3 showed that "wear" in English has a larger semantic category than "穿" in Chinese. It is possible to say "wear accessories" in English, but it is not acceptable to express that meaning using "穿" in Chinese.

Table 3

Collocation Information of "wear（穿）" in L1 and L2

	英語	中国語
coat 大衣		
jeans 牛仔裤		
skirt 裙子		**穿**
shoes 鞋		
socks 袜子		
hat, cap 帽子		
gloves 手套		
mask 口罩	**wear**	戴
earrings 耳环		
necklace 项链		
watch 手表		
rings 戒指		
glasses 眼镜		
necktie 领带		系，打
perfume 香水		抹，喷
smile 微笑		面带

Table 4 showed that "break" in English has a larger semantic category than words "打破，弄坏" which have similar meanings in Chinese. The expression of "break" in English can be not only "打破，弄坏" but also "打碎，违反，违背，破坏，分散" etc. in Chinese.

Table 4

Collocation Information of "break（打破，弄坏）" in L1 and L2

	英語	中国語
window 窗户窓		
record 记录		
tradition 传统		**打破**
deadlock 僵局		
silence 寂静		
camera 照相机		
TV 电视		**弄坏**
machine 机器		
cup 杯子		
glass 玻璃		打碎
vase 花瓶		
biscuit 饼干		掰开，弄碎
chocolate 巧克力		
rules 规定	**break**	
law 法律		违反
contract 合约		
agreement 协定		
promise 诺言		违背
serve 发球局		破
code 密码		破译
friendship 友谊		破坏
heart 心		使心碎
spirit 精神，意志		摧毁
concentration 注意力		分散
power 权力		削弱

3.4.3 Procedure

All participants were given the L2 association task and then the L1 association task. The stimulus words were ordered randomly while ensuring that the nouns related to the same verbs or the related verbs were not the items following. Two versions with inverse order were used. About half of the participants (N=18) used one version and the other half of them (N=19) used the other.

Participants were asked to read a written consent with the aim of the experiment and the private and confidential clause on it and sign it if he or she agrees to take part in the experiment and agrees the collocated data to be used in this study (Appendix B). All the participants sign the consent. To start off, they were told that the aim of the experiment was to investigate the state of lexical network in L1 and L2 mental lexicon. Then experiment began with oral instructions along with written instructions in L1, followed by the implementation of tasks. The order of conducting association tasks was the same for all participants, first L2 and then L1. Between L2 word association tasks and L1 word association tasks, a questionnaire was given to get information about their English learning background (Appendix A). Prior to the actual L2 word association tasks, an example was given to familiarize the participants with the way of association. A set of task materials with every stimulus word in the center of a circle were given and the participants' task was to write down around the circle as many associated words as possible for the stimulus word within two minutes (See Appendix A for more details).

3.4.4 Categorization of the Association Responses

The category criterion in this paper is referring to the category taxonomy and descriptions in Fitzpatrick (2006, 2007, 2009) and Fitzpatric and Izura (2011) and modified slightly. Each of the associated word responses obtained were classified into seven categories of association according to the lexical relationship between stimulus words and association responses: equivalent meaning, nonequivalent meaning, meaning and collocation, collocation, form and meaning, form and others. Response words were classified as others when there were responses which were just a copy of the stimulus words, or the responses are considered have no presumable semantic relation between the stimulus words and responses. Descriptions and some examples appeared in this study for each category are given in Table 5 below.

Table 5

Categorization of Association Responses

Category		Category	Description	Examples
Meaning-based	1	Equivalent meaning	Association responses which have equivalent meaning to the meaning of the stimulus word (mainly including synonym (defining synonym, specific synonym), coordination, super-ordination, lexical set, et al.	coat-skirt, coat-clothes
	2	Nonequivalent meaning	Association responses which have some kinds of conceptual links other than equivalent meaning to stimulus word.	medicine-hospital, soup-mother
	3	Meaning and collocation	Association responses which have both general meaning links to and a tendency to co-occur in the languages with stimulus words.	soup-vegetable, soup-bread
	4	Collocation only	Association responses which tend to co-occur in the languages with stimulus words but do not have any meaning relations.	medicine-box, wear-in
	5	Form and meaning	Association responses which have both general meaning links and form relations with stimulus words	medicine-medical, soup-sour, drink-drive
Form-based	6		Association responses which have form relations with stimulus words but do not have any meaning relations, including relationship of affix (inflectional affix and derivational affix) and similar orthographic form in only.	drink-drinks, break-brick
Others	7		Association responses which are not legible, just the same as the stimulus words, or have no decipherable links to the stimulus words.	drink-drink, drink-peer

Association English words, which are considered as spell miss and the real words of them are

apparent, are corrected by the writer and counted as the word it should be.

3.4.5 Statistical Analysis

The associated words of the randomly selected stimulus "medicine, wear, drink" ((1/3 of the L2 data), and "打破，喝，大衣，穿" (2/5 of the L1 data) of all thirty-five participants' were classified and labeled as 1 to 7 by the author and another person (one of the participants in this experiment) based on this classification standard in according to Table 5. Because it is the first time for the participant to classify this kind of words, before doing independently, the associated word data of the randomly extracted stimulus words "break" and "药" were used as a classification training which conducted according to the criteria by the author and the participant together. The raters' consistency was 92% (1,508 of 1,636 items) and it is an enough high inter-rater reliability according to Shweta, Bajpai, and Chaturvedi, 2015. Therefore, the author classified the remaining items. One of the participants showed an abnormally high percentage of responses coded 7, so the data was excluded from further analysis. Thus, the valid data become thirty-four (nineteen females, fourteen males and one without information) lastly. In other words, there was a total number of 4432 association responses from thirty-four participants, of which 1943 comes from L2 English, 2489 comes from L1 Chinese (see Table 8). This indicates that respondents were familiar with most stimulus words. As for there were nine English but ten Chinese stimulus words, the average of "打破" and "弄坏" in L1 which are designed as the equivalents of "break" in L2 was counted and used, in convenience to the comparison of L1 and L2 about numbers. The average was used in both total data and individual data for analysis.

3.5 Results

Two word association data sets were produced for each participant: responses for the ten L1 stimulus words and nine L2 stimulus words.

3.5.1 Comparing the L1 and L2: Number of Association Responses per Category

The total number of L2 association responses was 1943 and the total number of L1 association responses was 2285 (after adjustment, as explained above, a conversion of number by averaging the data of two particular words among them was used to ensure an equivalence

between L1 and L2). The general pattern is that most of the reported associations are, in this way or that way, meaning related (around 95%). Few associations are form-based or belong to "Others" (around 5%). In general, participants gave fewer responses to the L2 stimulus words than to the L1, however the paired t-test showed that there was no significant difference between the number of association numbers of L1 and L2 (t (33), 1.857, p= .072). Then the number of association responses in L1 and L2 was submitted to a correlation analysis. There was a correlation of the number of association responses between L1 an L2 (r= .717, t (32) = 5.811, p< .001).

To know more details about the association responses in L1 an L2 and compare response category behavior across the L1 and L2, first, the number of responses per participant in each category was counted (Table 6). The number of association responses of each category in L1 and L2 was submitted to a correlation analysis. In the result, there was a correlation between L1 and L2 in Category 1, Category 2, category 3, and Category 7. This means there was a roughly similar condition of association response in L1 and L2.

Table 6

The Results of Correlation Analyses of Association Responses' Number Between L1 and L2 for Each Category

Category	r	t (32)	*p*
Category 1	.719	5.845	.000
Category 2	.727	5.983	.000
Category 3	.685	5.320	.000
Category 4	-.044	-0.247	.807
Category 5	-.064	-0.362	.720
Category 6	-.010	-0.059	.954
Category 7	.693	5.441	.000

The number of association responses was submitted to a 2 × 7 ANOVA with category of association response and language as within-subject factors. Results revealed that the interaction between language and category was significant, F (6, 198) = 8.406, MSe = 24.505, p< .001, η

$_p^2$= .203. The main effect of category of association response was also significant, F (6, 198) = 72.508, MSe = 116.141, p< .001, η_p^2= .687, with some categories producing more responses than others. However, the main effect of language was not significant, F (1, 33) = 3.449, MSe = 71.252, p> .05, η_p^2= .095. The analysis of simple main effect of language to category shows a significant effect in Category 1 (equivalent meaning), d= .437, p< .001 and Category 3 (meaning and collocation) d= .520, p< .001 (See *Figure* 1).

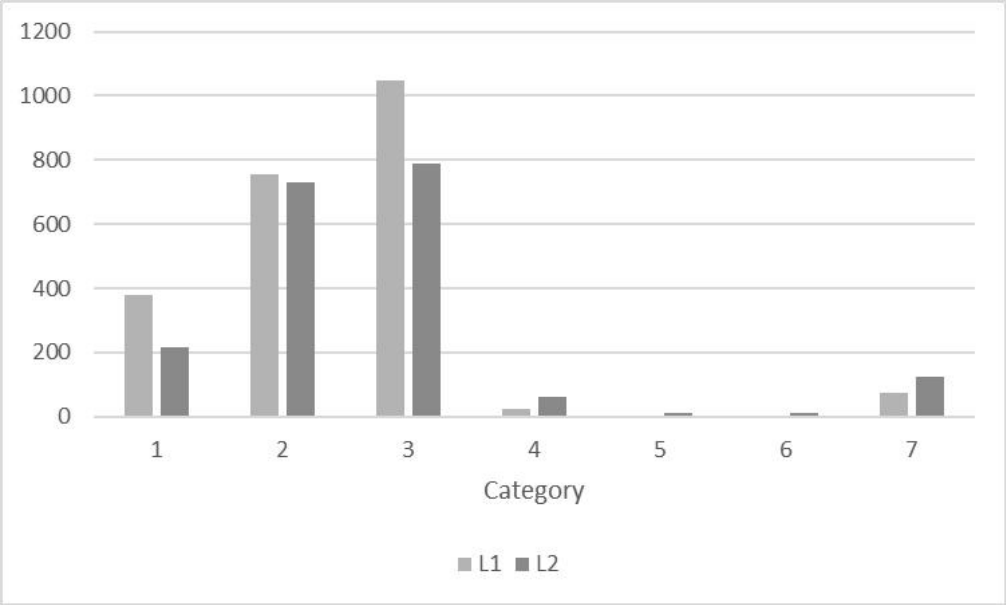


Figure 1.

Number of association responses by category.

The analysis of simple main effect of category to language shows a significant effect in both L1 and L2 (p< .001), which means that some categories producing more responses than others in each language (See *Figure* 2 and Table 7).

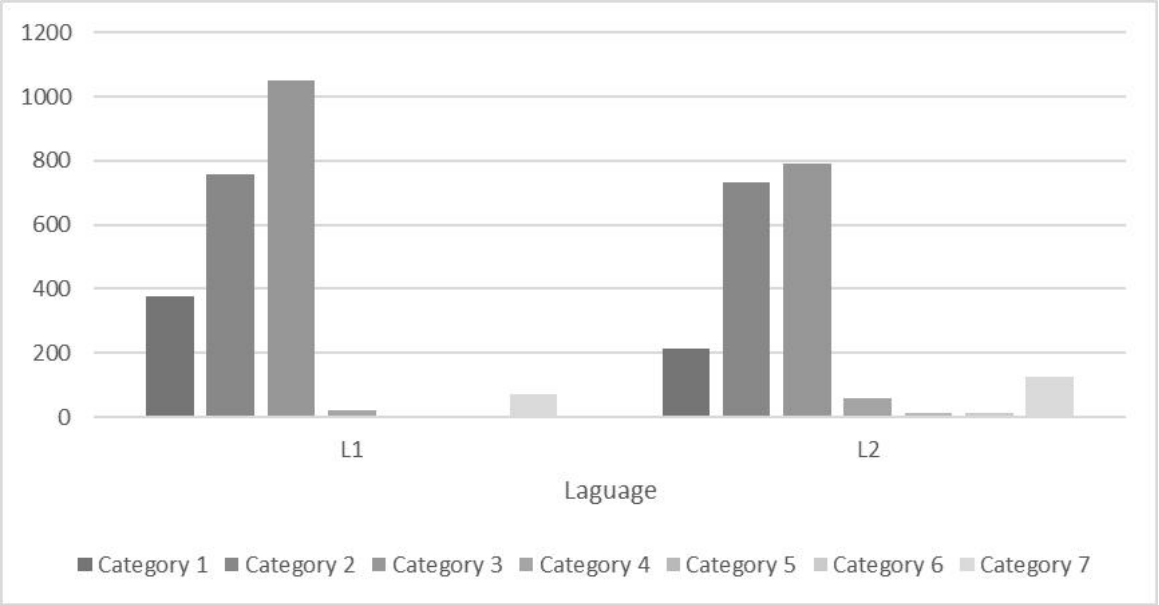


Figure 2.

Number of association responses by language.

Category 2 (nonequivalent meaning) and Category 3 (meaning and collocation), are the most frequently produced two kinds of responses. The response number of the two categories are significantly more than any other categories in both L1 and L2. *Figure* 2 clearly exhibits the significantly larger number of Category 2 (nonequivalent meaning) and Category 3 (meaning and collocation) than all other categories in both L1 and L2. Table 7 shows more details about the category difference.

Table 7

Simple Main Effect of Category to Language

	L1		L2	
Category	d	*p*	d	*p*
1 - 2	-.580	ns	-1.077	.003
1 - 3	-1.359	.000	-1.159	.000
1 - 4	1.422	.000	.614	ns
1 - 5	1.506	.000	.813	.010
1 - 6	1.510	.000	.813	.011
1 - 7	1.086	.000	.320	ns
2 - 3	-.399	ns	-.072	ns
2 - 4	1.217	.000	1.118	.000
2 - 5	1.248	.000	1.199	.000
2 - 6	1.250	.000	1.199	.000
2 - 7	1.107	.000	.990	.000
3 - 4	2.396	.000	1.694	.000
3 - 5	2.444	.000	1.811	.000
3 - 6	2.447	.000	1.811	.000
3 - 7	2.183	.000	1.481	.000
4 - 5	.468	ns	1.284	.000
4 - 6	.495	ns	1.285	.000
4 - 7	-.377	ns	-.469	ns
5 - 6	.062	ns	.000	ns
5 - 7	-.522	ns	-.847	.011
6 - 7	-.529	ns	-.847	.010

3.5.2 Comparing the L1 and L2: Number of Association Responses in the First to Twentieth Associations

Figure 3 below represents the number of association responses in the first to twentieth of L1 and L2.

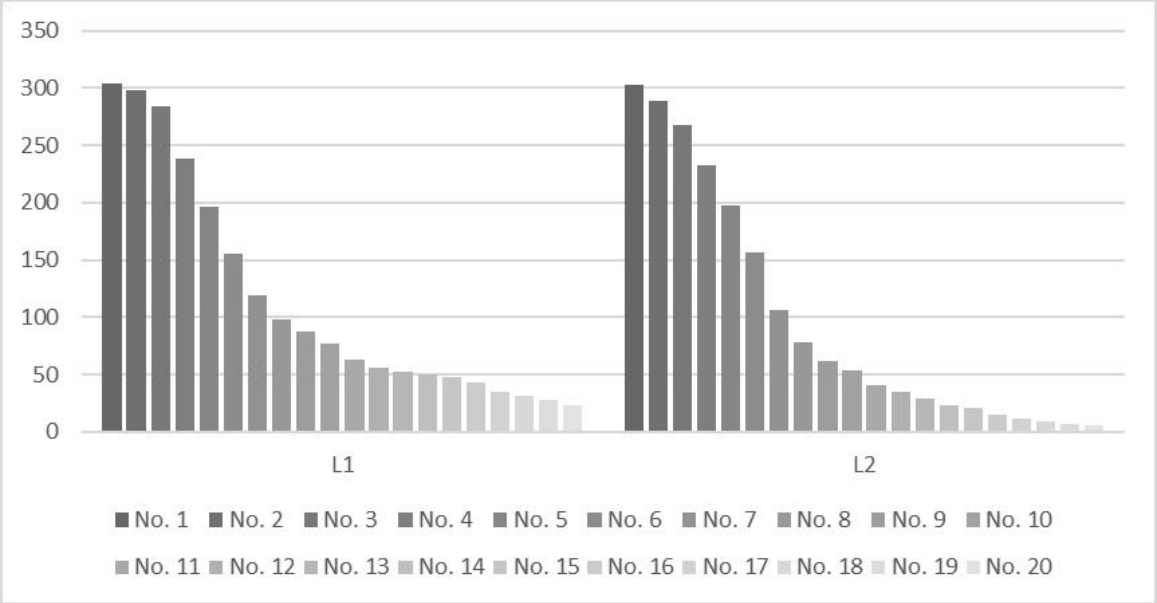


Figure 3.

Number of association responses from the first to the twentieth by language.

The results suggest that both for L1 Chinese and L2 English associations, there is an apparent and similar tendency of decrease of the number of association responses from the first association to the twentieth.

The average number of association responses for each word in L2 was about six and for each word in L1 was about seven. Therefore, the number of association responses of each category in the first to seven of L1 and L2 is represents in the *Figure* 4 below to show more details.

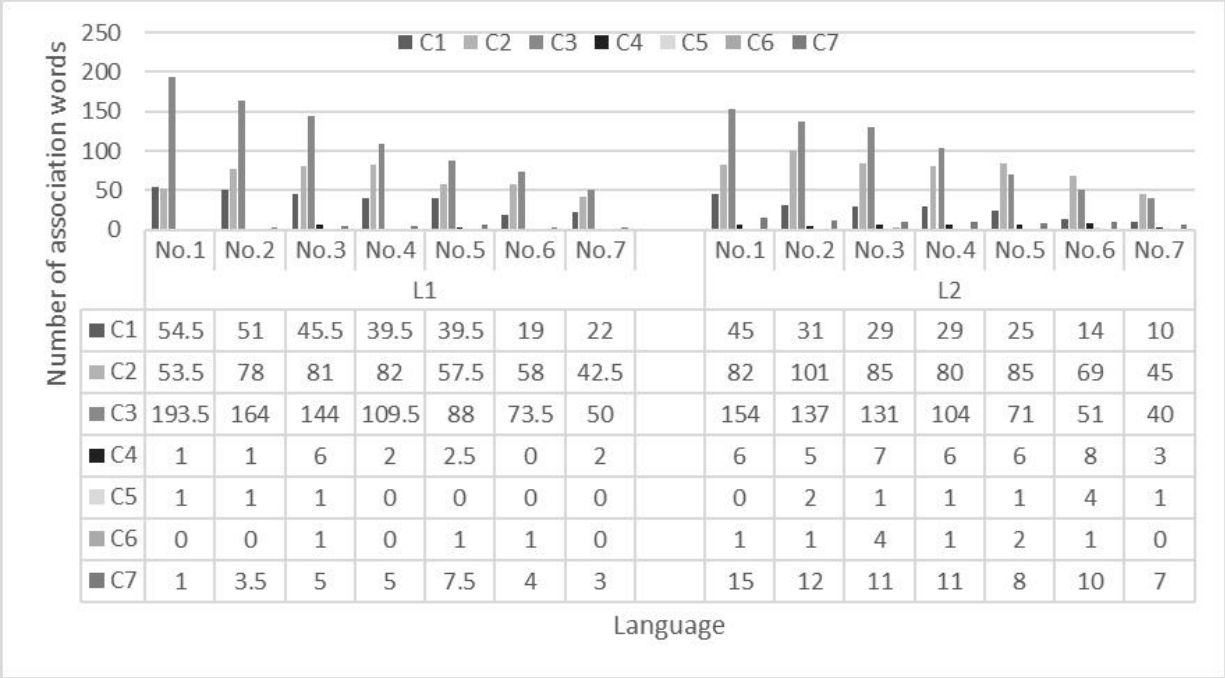


Figure 4.

Number of association responses in the first to seventh by category.

There is a clear tendency that the most frequent associations are belong to Category 3 in L1, and then followed by Category 2, except the first association words produced which followed by category 1. However, for the first to fourth association words in L2, the most frequent associations are belonging to Category 3, while for the fifth to seventh words, the most frequent association words are belonging to Category 2. The least frequent association words are words belong to Category 5 and Category 6 in both L1 and L2.

3.5.3 Comparing the Number and Category of Association Responses per Word

The total number of association responses produced by all the participants for each word per category is summarized in Table 8 below.

Table 8

Number of Association Responses per Word and per Category

category		1. Equivalent meaning	2. Nonequivalent meaning	3. Meaning and collocation	4. Collocation only	5. Form and meaning	6. Form only	7. Others	Total
drink		3	50	117	14	3	8	15	210
wear		5	57	144	1	0	0	8	215
break	(verb)	5	23	28	15	0	2	45	118
	(noun)	1	80	2	4	1	1	0	89
medicine		8	163	60	9	5	0	13	258
soup		28	57	116	12	1	1	9	224
porridge		48	52	79	4	2	0	14	199
coat		34	77	92	1	0	0	4	208
jeans		46	66	67	0	0	0	6	185
shoes		36	107	84	0	0	0	10	237
English (L2) in all		214	732	789	60	12	12	124	1943
喝		10	76	168	2	0	0	12	268
穿		21	43	144	7	0	0	24	239
打破、弄坏	打破	19	50	110	3	0	0	21	
	弄坏	27	77	112	0	0	0	7	213
	平均	23	63.5	111	1.5	0	0	14	
药		32	151	91	6	0	0	1	281
汤		39	72	138	0	0	0	4	253
粥		74	90	107	0	0	0	2	273
大衣		40	94	105	0	0	0	4	243
牛仔裤		56	53	95	5	4	3	4	220
鞋		83	113	91	0	0	0	8	295
Chinese (L1) in all		378	755.5	1050	21.5	4	3	73	2285
Total		592	1487.5	1839	81.5	16	15	197	4228
L1-L2 (difference)		164	23.5	261	-38.5	-8	-9	-51	342

The following *Figure* 5 showed the detailed category condition of the number of association responses produced for each stimulus word.

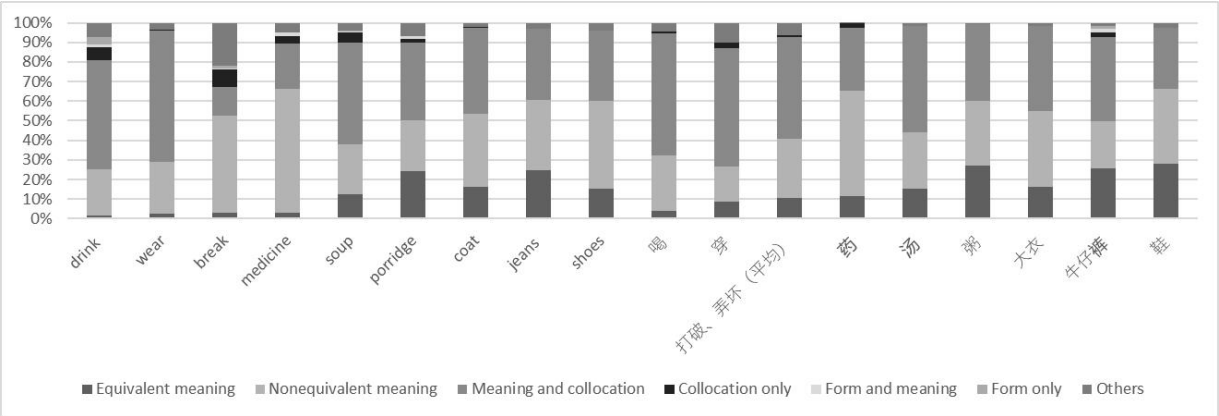


Figure 5.

Category of association responses for each stimulus word.

The number of association responses of equivalent meaning words in L1 and L2 was submitted to a correlation analysis (Table 9). In the result, there was a correlation between all word pairs with equivalent meaning in L1 and L2. This means there was a roughly similar condition of association response in L1 and L2.

Table 9

Correlation Analysis's Results of Equivalent Stimulus Words Between L1 and L2

	r	t (32)	*p*
drink 喝	.768	6.778	.000
wear 穿	.729	6.021	.000
break 打破 弄坏	.463	2.952	.006
medicine 药	.717	5.814	.000
soup 汤	.624	4.523	.000
porridge 粥	.545	3.678	.001
coat 大衣	.598	4.220	.000
jeans 牛仔裤	.636	4.657	.000
shoes 鞋	.660	4.976	.000

3.5.4 Comparing the Verbs and Nouns: Number of Association Responses per Category

The number of responses for stimulus nouns and stimulus verbs are showed in *Figure* 6 (L1) and *Figure* 7 (L2) below. As for the word "break" might have been considered as both a verb and a noun in English by some participants from the association responses they gave, the results of L2 just take another two words "drink" and "wear" as verbs. In L1, both the results of just two verbs "喝" and "穿" and taking all four verbs (喝，穿，打破，弄坏) into account were given. We made a make a distinction between Verb (2) and Verb (4) in *Figure 6.*

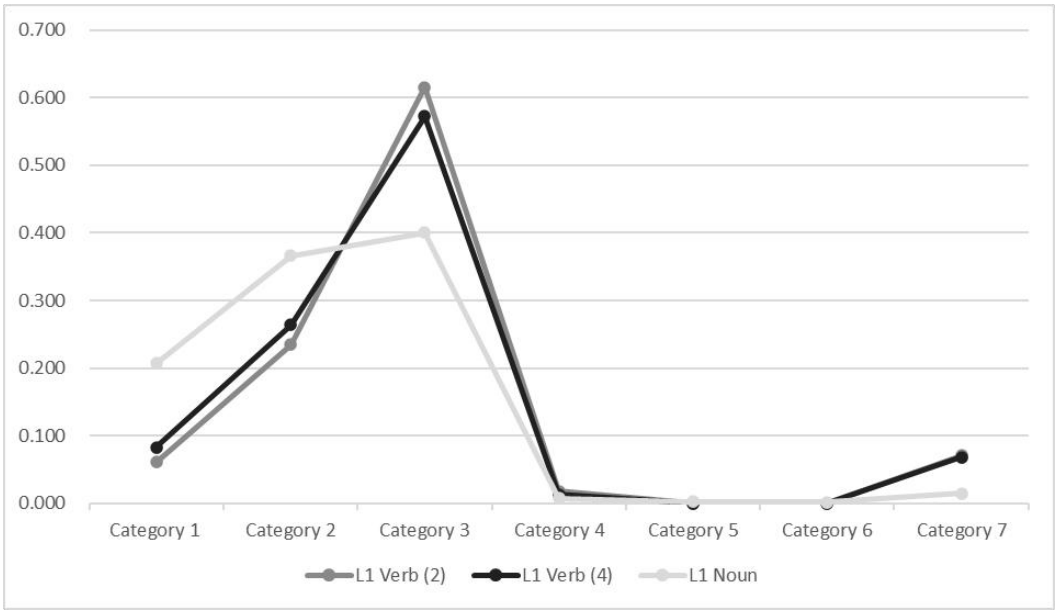


Figure 6.

Comparison of stimulus nouns and verbs in L1.

The most frequent association response was Category 3, then following Category 2 and Category 1 for both L1 verbs and nouns, while the percentage of Category 3 and the difference of percentage between Category 3 and Category 2 in verbs was higher and larger than in nouns.

Figure 7 below showed the number of responses for stimulus nouns and stimulus verbs in L2.

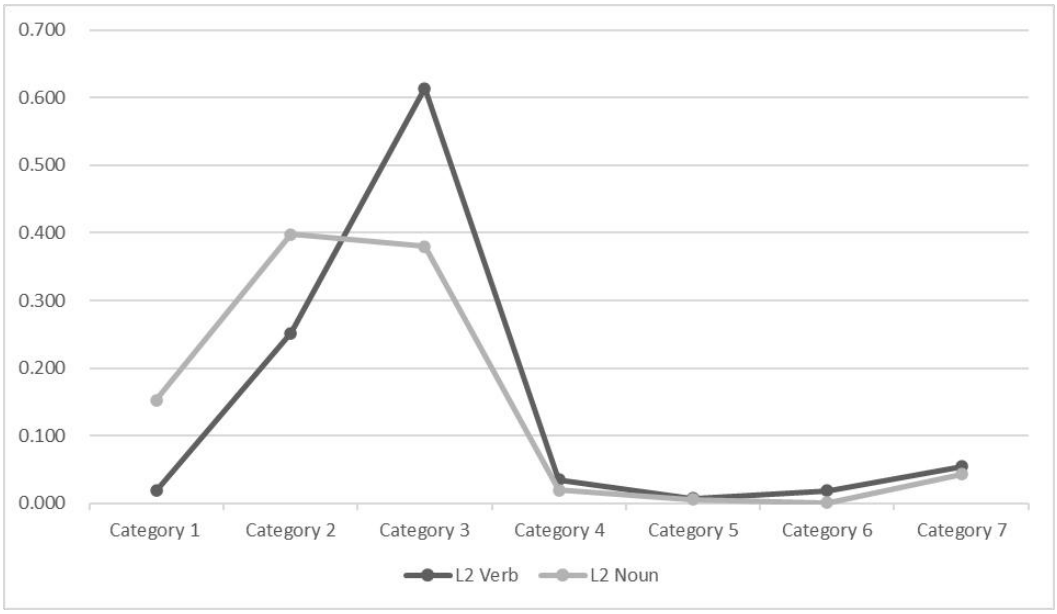


Figure 7.

Comparison of stimulus nouns and verbs in L2.

If we do not consider Category 7, the most frequent association response was Category 3, then following Category 2 and Category 4 and Category 1 for L2 verbs, while the most frequent association response was Category 2, then following Category 3 and Category 1 for and nouns. Moreover, the percentage of Category 3 and the difference of percentage between Category 3 and Category 2 in verbs was higher and larger than in nouns.

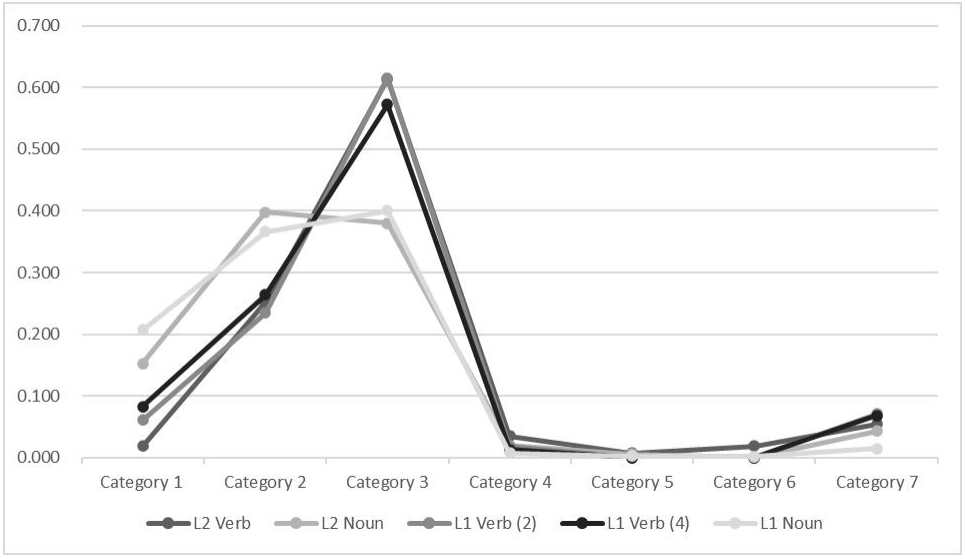


Figure 8.

Comparison of stimulus nouns and verbs in L1 and L2.

Figure 8 put both results of L1 and L2 together which reveals a similar general tendency of the difference between stimulus nouns and verbs across language but shows some intricate difference between them. There was a larger portion of responses in meaning and collocation category for stimulus verbs than nouns, but a larger portion of responses in equivalent meaning and nonequivalent meaning for stimulus nouns than verbs.

To sum up, Table 10 presents the ratio of association responses per category for verbs and nouns, respectively.

Table 10

The Ratio of Association Responses per Category for Verbs and Nouns, Respectively

Category	L1 Verb (2)	L1 Verb (4)	L2 Verb	L1 Noun	L2 Noun
1	6.1	8.3	1.9	20.7	15.3
2	23.5	26.4	25.2	36.6	39.8
3	61.5	57.2	61.4	40.1	38.0
4	1.8	1.3	3.5	0.7	2.0
5	0.0	0.0	0.7	0.3	0.6
6	0.0	0.0	1.9	0.2	0.1
7	7.1	6.9	5.4	1.5	4.3

In both L1 and L2, as expected, the differences in association types depending on the part of speech were observed. There was no significant difference in the ratio of association types between L1 and L2. When the stimulus word is a verb, the proportions of Category 3: Meaning and collocation and Category 2: Nonequivalent meaning are high, accounting for about 85% totally, and about 60% for Category 3. On the other hand, when the stimulus word is a noun, in addition to Category 3: Meaning and collocation and Category 2: Nonequivalent meaning, the ratio of Category 1: Equivalent meaning is also high, accounting for about 90% of the total, 80% of Category 3 and Category 2, and about 15-20% for Category 1.

From a different angle, Category 2: Nonequivalent meaning and Category 3: Meaning and collocation ratios were high for both verbs and nouns, but Category 3 accounted overwhelmingly high for about 60% of in the situation of verb stimulus words. However, in the case of nouns, the ratios of Category 3 and Category 2 were almost similar, about 40% each. In other words, the difference is that the ratio of Category 3 was about 20% higher for verbs, and the ratio of Category

1: Equivalent meaning appeared about 20% for nouns.

3.5.5 Qualitative Analysis of Association Responses for Words with Different Relationship Between the Semantic Categories of L1 and L2

The result of word association was actually different according to the relationship between the semantic categories of L1 and L2. For stimulus words “medicine”. “soup” and “porridge,” the association verbs which compose proper collocations in L1 but improper collocations in L2 like “drink” or “eat” actually appeared in L2 English word association. The verb “drink” appeared once and “eat” appeared five times as association responses among all 258 responses for stimulus words “medicine,” while the verb “take” which is considered as proper collocation in L2 just appeared once. The verb “drink” appeared ten times as association responses among all 224 responses for stimulus words “soup,” while the verb “eat” which is considered as proper collocation in L2 just appeared once. The verb “drink” appeared once and “eat” appeared ten times as association responses among all 199 responses for stimulus words “porridge,” while the verb “eat” which is considered as proper collocation in L2 just appeared five times. Moreover, for the stimulus word “drink”, nouns which composes proper collocations with “喝” in L1 but improper collocations with “drink” in L2 appeared. For example, “soup” (5/210) and “porridge” (1/210) appeared as the association responses in English word association (see Table 11 for more details).

Table 11

The Association Results of “drink, medicine, soup, porridge”

Stimulus word	Number of association words	Association from L1transfer (times)	Correct L2 expression
medicine	258	drink (1), eat (5)	take (1)
soup	224	drink (10)	eat (1)
porridge	199	drink (10)	eat (5)
drink	210	soup (5), porridge (1)	

As showed in Table 12 below, on the other hand, in all 215 association responses for the

stimulus word “wear”, besides “coat”, “jeans” and “shoes,” which compose collocations with “wear” in L2 just same as “大衣”, “牛仔裤” and “鞋,” which compose collocations with “穿” in L1, some other association words compose collocations with words like “戴” or “带着 （面带）,” instead of “穿” in L1, and are not acceptable as collocations with “穿” but all compose collocations with “wear” in L2 actually appeared. For example, there were association words including nouns like “cap” (2), “glasses” (8), “glove” (or “gloves”, 3, totally), “hat” (or “hats”, 13, totally), “jewelry” (3), “mask” (1), “rings” (1), “scarf” (1), and “smile” (1).

Similarly, for the stimulus word “break” in L2, the association responses included association responses of equivalent stimulus words “打破,” “弄坏,” “打碎,” “破碎,” “破裂” and so on in L1. For example, words like “cup” (5), “friendship” (1), “glass” (or “glasses”, 3), “heart” (1), “records” (1), “rule” (or “rules”, 6), “vase” (1) and “window” (4) appeared in the association responses in L2 word association. It is worth noting that the stimulus word “break” is both a verb and a noun, some participants gave association responses based on considering “break” as a verb, some gave association responses based on considering “break” as a noun, and others gave association responses considering it as both a verb and a noun. Therefore, the association responses for the stimulus word “break” became complicated beyond hypothesized.

Table 12

The Association Results of “drink, medicine, soup, porridge”

Stimulus word	Number of association words	Correct L2 expression with no L1 homologous expression (times)
wear	215	cap (2), glasses (8), glove(s) (3), hat(s) (13), jewelry (3), mask (1), rings (1), scarf (1), smile (1)
break	118 (as a verb) 89 (as a noun)	cup (5), friendship (1), heart (1), records (1), rule(s) (6), vase (1), window (4)

3.6 Discussion

Discussions were conducted focusing on the research questions of present study one by one.

Firstly, about Research Questions 1: There was a correlation but not significant difference between the number of association responses in L1 and L2 in general. Correlation analyses

showed a correlation between L1 and L2 in Category 1, Category 2, category 3, and Category 7. The significant interaction between language and category, significant simple main effect of language to category in Category 1 (equivalent meaning) and Category 3 (meaning and collocation), and significant simple main effect of category to language in both L1 and L2, means that the number of responses regarding to category (Category 1 and Category 3) are different according to language, and the number of responses regarding to L1 and L2 are varied according to category. Responses belong to Category 2 (nonequivalent meaning) and Category 3 (meaning and collocation) were more than others in both L1 and L2. These results were different from the results of Fitzpatrick and Izura (2011) in several points, first, the most frequent response category was nonequivalent meaning in their study, but not only nonequivalent meaning but also meaning and collocation in this study. Then, significant differences were found in the equivalent meaning category in Fitzpatrick and Izura (2011), but in not only the equivalent meaning category but also meaning and collocation category in this study, with more responses produced in the L1 than L2. The high frequency of nonequivalent meaning category of both languages supports Fitzpatrick and Izura (2011)'s idea that links based on form and meaning do not always exist or most stimulus words would have been had links with a large number of lexical items that would fit this nonequivalent meaning category essentially. The significantly less equivalent meaning category in L2 than in L1 supports research results in many previous studies that paradigmatic connection is considered more difficult to be constructed. The significantly more meaning and collocation category in L1 and the high frequency in both languages of meaning and collocation can be a support to Fitzpatrick and Izura (2011)'s explanation that there are several different kinds of potential activation triggers (meaning, collocation, form) and that the activation event is stronger when more than one of these is present. Moreover, the strength of each activation trigger may be stronger in L1 than in L2. The high frequency of meaning and collocation and a significant difference between L1 and L2 regarding to this category may suggest that this "dual-link response" mentioned in Fitzpatrick and Izura (2011) was activated easier in L1 than in L2.

Next, about Research Questions 2: There was an apparent and similar tendency of decrease of the number of association responses from the first association to the later ones in both L2 and L1 generally. The similarity in the most frequent association response category for the first to fourth association words and difference in the most frequent association response category for the fifth to

seventh words between L1 and L2, also reveals that there are more meaning and collocation category response in L1 than in L2.

Then, about Research Questions 3: The stimulus words' part of speech (verbs or nouns) actually showed different response results. The result of a larger portion of Category 1: Equivalent meaning responses of nouns supports the results in previous studies (Cronin, 2002; Nissen & Henriksen, 2006) that nouns elicit paradigmatic responses more often than verbs do. The explanation can be considered from the following aspects. First, nouns may have been had connections with a large number of lexical items that would fit this equivalent meaning category than verbs, essentially. Then, as with more concrete semantic features, larger number in language which may lead to higher chance of encounter, or possibility of earlier encounter for language users, therefore, nouns are generally considered to be more highly integrated into the mental lexicon than verbs. In addition, nouns tended to lead to a considerable proportion of nonequivalent meaning associations, yielding responses such as "grandmother" or "mother" for the stimulus word "soup", which is related to the participants' own life experience. The result of a larger portion of responses in meaning and collocation category for stimulus verbs than nouns and a larger portion of responses in nonequivalent meaning for stimulus nouns than verbs may also reveal the higher portion of syntagmatic-related connections in verbs.

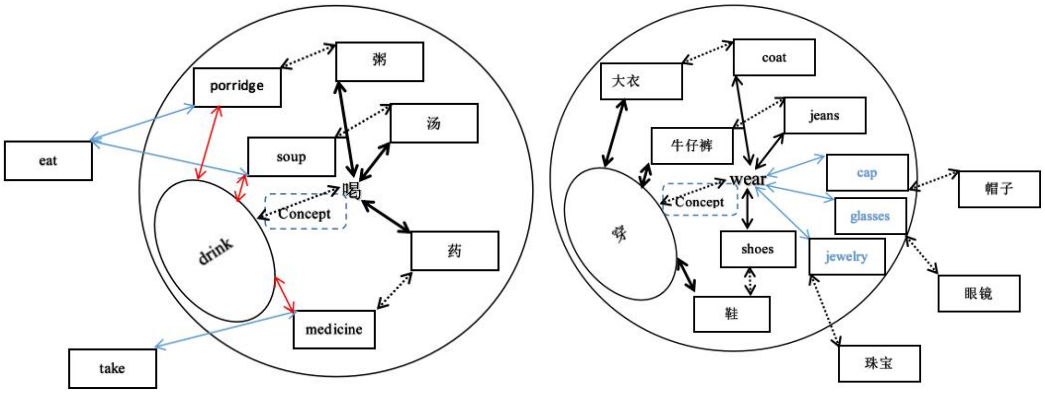
Even we didn't have adjectives in this study, on one hand, adjectives were usually considered similar to nouns to yield more paradigmatic associations. Adjectives relate to each other via meanings, antonyms, or synonyms. However, on the other hand, adjectives were also prone to syntagmatic associations, resulting in "adjective + nouns" pattern. This can be explained by the fact that most adjectives are abstract in nature and require "a noun for a fully-fledged interpretation since they are semantically under specified" (Li & Wang, 2016, p. 454) if in isolation. Moreover, the adjectives and verbs are easy to prompt syntagmatic responses than nouns may because both the L1 Chinese and L2 English are SVO languages. The language syntax order is "adjective + nouns" and "verbs + nouns", so it is easier for participants to associate nouns from adjectives or verb than from nouns to verbs or adjectives.

In a word, the similar tendency in both L1 and L2 states a common difference caused by words' part of speech which is independent of language.

Lastly, about Research Questions 4: The results of L2 word association were different with

the different relationship between the semantic categories of L1 and L2. The difference between association words concerning “drink” (“medicine”, “soup”, “porridge”) which in L1 (Chinese) have a broader semantic category and “wear” (“coat”, “jeans”, “shoes”) which in L2 (English) have a broader semantic category verified Bi (2020)’s proposition that words like “drink” which in L1 (Chinese) have a broader semantic category are considered to be more difficult to acquire than words like “wear” and “break” which in L2 (English) have a broader semantic category. These results confirmed the mediation of L1 word knowledge in construction of L2 word network, and indicate the influence of L1 word knowledge to word association was different with the different relationship between the semantic categories of L1 and L2.

According to the results above, in *Figure* 9, the network of L1 and L2 taking (a) drink‧喝 (medicine, soup, porridge) and (b) wear‧穿 (coat, jeans, shoes) as examples are showed. The size of circle means the range of semantic category of L1 and L2 (L1 “喝” > L2 “drink” in Figure 9 a; L1 “穿” < L2 “wear” in Figure 9 b). The overlapping relations of semantic category between L1 and L2 was not deliberated in these figures. In other words, the overlapping parts of circles in these figures may be imprecise and the range of semantic category was the only implication of the circles. The intralinguistic links are showed by solid lines, the interlinguistic links are showed by dotted lines. Some inappropriate associations which negative evidence is considered necessary are showed in red, and which positive evidence necessary are showed in blue. In *Figure* 9 (b), some words whose collocation is appropriate for “wear” but inappropriate for “穿” were also in blue.



a b

Figure 9.

L1 and L2 network image (a. drink‧喝; b. wear‧穿).

3.7 Conclusion

3.7.1 Summary of the Findings

The present study investigated features of Chinese EFL learners' word networks in both L1 and L2 by analyzing word association responses. The present study also tried to investigate how it is different or similar in word association responses between L1 and L2, between the association responses produced earlier and later, and between stimulus words with different part of speech. Furthermore, the present research focused on whether the results of word association differ according to the relationship between the semantic categories of L1 and L2. In other words, are there some differences between the results of words like "drink" which in L1 (Chinese) have a broader semantic category, and the results of words like "wear" and "break" which in L2 (English) have a broader semantic category. Two word association tasks (English and Chinese) were implemented to Chinese university students and the association responses elicited were categorized into seven categories (equivalent meaning, nonequivalent meaning, meaning and collocation, collocation, form and meaning, form and others) according to the lexical relationship between stimulus words and association responses. This study found that there was a correlation but not significant difference between the number of association responses in L1 and L2 in general. There was an interaction between language and category to the number of association responses, significant simple main effect of language to category in equivalent meaning category and meaning and collocation category, and significant simple main effect of category to language in both L1 and L2. This study also found a decrease tendency of the number of association responses from the first association to the later ones in both L1 and L2. Moreover, some differences of response category elicited by stimulus words' part of speech (nouns or verbs). In addition, the difference between association of words concerning "drink" (including "medicine", "soup," and "porridge"), which in L1 (Chinese) have a broader semantic category and "wear" (including "coat", "jeans," and "shoes"), which in L2 (English) have a broader semantic category appeared. The results of this study suggest that the category type and number of word association responses varied with languages, stimulus words' part of speech, and sequences the responses were elicited in intricate ways. This study also confirmed the mediation of L1 word knowledge in construction of L2 word network and suggested that the influence of L1 word knowledge to L2 word association is different with the different relationship between the semantic categories of L1 and

L2. These findings were discussed together with the construction of word network in mental lexicon and the process of L2 vocabulary acquisition. Although the findings support Fitzpatrick and Izura (2011)'s propose of "dual-link responses" in general, there remain questions about the nature of both intralingual and interlingual networks during the word acquisition process, and about the exact way of activation of specific triggers in the output process.

3.7.2 Pedagogical Implications

As for SLA, from the results and discussions above, words association features varied with language and words' part of speech which reveals that in fact lexical networks are different in some way, moreover, L2 words' network may be different according to the relationships between the L1 and L2 semantic category. This suggests that different types of vocabulary may need to be taught in different ways. Teachers need to know the semantic feature and even the semantic relationships between L1 and L2 of each word and give support accordingly.

For example, when teaching vocabulary like "wear" and "break," that have an L2 semantic category broader than L1, only positive evidence is enough, so teachers should offer positive evidence as much as possible. However, when it comes to vocabulary with an L2 semantic category narrower than L1, like "drink," L1 transfer is used and the risk of over-generalization is great, so teachers should provide more explicit negative evidence and be patient with the learners' development.

Besides, teachers need also pay more attention to the words' part of speech as there may have differences among nouns, verbs, and adjectives in the word network construction process. Therefore, to adopt corresponding teaching or learning approach for different part of speech's words may help reinforce specific links in learners' network and facilitate the process of vocabulary acquisition.

3.7.3 Limitations and Recommendations for Further Research

Some limitations of this study and recommendations for further research are summarized.

First and foremost, although the categorization of the associations has improved a lot, there is still not a consistent criterion agreed by most researchers. Some response words are still difficult to be classified precisely even refer to the recent standard of Fitzpatrick and Izura (2011) which

was referred in this study. Moreover, as the category taxonomies were mainly made for English word association tasks, the form-based connection for the alphabetical language may not always be appropriate for other languages like Chinese or Japanese, which are considered as character languages. Therefore, the category taxonomy for word associations still in need of improving in the future.

Next, the influence of stimulus words was discussed in many studies so far. As a special purpose to investigate the difference between association of words which in L1 (Chinese) have a broader semantic category and those which in L2 (English) have a broader semantic category, the semantic category of words was considered firstly, rather than word frequency, words' part of speech, word length, and so on. Even the not large number of responses belonging to category Others could explain the familiarity of these stimulus words to some extent, more exact assurance of stimulus words' quality is needed in future research. In addition, the neglect of the dual part of speech of word "break" in the period of stimulus word selection, and the different number of stimulus words in L1 and L2 association tasks made some troubles for analysis. Further research is expected to be implemented after sweeping away these obstacles.

Then, about the rater, the judgment of whether a pair of stimulus words and association responses belongs to a specific category, collocation category, for example, in a foreign language is difficult for a L2 learners sometimes. This calls a necessary for native raters but it is not that easy to put into practice for all research in fact. Furthermore, even there are native raters, the raters' idea about a word may not always same as the test takers. Thus, an approach which participants themselves coding their own responses according to a specific category criterion after word association tasks by themselves (similar to the retrospective interview in Fitzpatrick, 2006) may be useful to some extent, although there will also be some practical difficulties. Using of both test takers' self-coding and test implementers' coding should be newly attempted in future research.

In addition, two word association tasks in this study were implemented successively as a L2 task followed by a L1 task with only a questionnaire about English learning background between them. In further research, a period between the two word association task and the sequence of L1 and L2 word association may also need to be considered to minimize the effect of each other.

Last but not least, cooperation of larger number of participants are expected and further

consideration of L2 proficiency as an additional influence factor may be important in future research.

CHAPTER 4 THE PRESENT STUDY: EXPERIMENT 2

Effects of Breadth and Depth of Vocabulary Knowledge on Word Association in L2 for Chinese EFL Learners

As mentioned in **2.3.2**, some researchers (Fitzpatrick, 2006, 2007; Nissen & Henriksen, 2006; Wolter, 2001; Namei, 2004) have argued that the participant's knowledge of stimulus words, especially the degree of vocabulary knowledge of each individual word, instead of the general language proficiency of the learner may vary types of association words. To get a better grasp of L2 mental lexicon, which further lead to more scientific views of L2 vocabulary learning and teaching, in Experiment 2, the effects of breadth and depth of vocabulary knowledge on word association in L2 for Chinese EFL learners were researched.

4.1 Research Questions

The research questions of present study were posited as follows:

Research Question 1: Does the vocabulary size (the breadth of vocabulary knowledge) affect Chinese EFL learners' word association responses?

Research Question 2: Are there significant differences of the number and the type of association responses in L2 for participants groups of different vocabulary size (HIGH and LOW)?

Research Question 3: Does the depth vocabulary knowledge affect Chinese EFL learners' word association responses? Is the number of association responses in L2 increases as the depth of knowledge of stimulus words increasing? Do response types change as well?

Research Question 4: Does the stimulus words, especially their part of speech (nouns, verbs, or adjectives) affect the response types?

4.2 Hypotheses

According to the previous studies, L2 word knowledge (both the size of vocabulary and how well a word is known) are key components for determining the integration degree of individual words that make up the structure of L2 mental lexicon.

Hypothesis 1: The vocabulary size (the breadth of vocabulary knowledge) affects Chinese

EFL learners' word association responses. We assume that there will be a correlation between vocabulary size and the number of word association responses for the L2 learners. The number of word association responses and the variety will develop as the vocabulary size increases.

Hypothesis 2: It is assumed that there will be a significant difference of the number of association responses in L2 for participants of HIGH vocabulary size and LOW vocabulary size and the type of association responses is assumed to show different characteristics.

Hypothesis 3: The vocabulary knowledge of the stimulus word is assumed to be an important factor that influences the word association types. The meaning-based association increases, while form-based association decreases along with the development of the depth of word knowledge.

Hypothesis 4: As many researchers found an influence of word's part of speech on response types in many previous studies and it has also verified in Experiment 1, we assume that the stimulus words' part of speech (verbs or nouns) will affect the response types and nouns will elicit paradigmatic responses more than adjectives or verbs do the same as in Experiment 1.

4.3 The Main Study

4.3.1 Participants

Forty Chinese EFL learners (thirty-one females, eight males and one without information) took part in the study. Four of them are university students (three students are at the third year and one is at the fourth year) and others have already graduated from universities and entered graduate schools as a master or doctoral student or are at work. All participants with Chinese Mandarin as their mother tongue have learned English as a foreign language for 13.44 years in average. Participants were an average of 25.6 years old (range: 20—32). None of them have an experience of going abroad to an English country for a long period.

4.3.2 Materials

The total vocabulary size (i.e., the breadth of word knowledge, how many words someone knows) was measured by the Vocabulary Size Test (VST) (Nation and Beglar, 2007). In an attempt to reach a more systematic approximation of learners' total vocabulary size, the VST samples words from sequential frequency bands from k1 to k14. Ten items were selected from each frequency band of 1,000 word families to compose the test. Hence, the sampled 140 items in the

test attempts to measure word knowledge of 14,000 word families. A newer version including items of frequency bands from k1 to k20 exists as well, but the original k1 to k14 version is the more widely used one. The word items are presented in non-defining minimal context sentences, followed by four definition options from which the learner needs to choose the correct definition of the word. There is monolingual version which with four definition options in English and bilingual versions which with four definition options in different languages as Arabic, Gujarati, Korean, Japanese, Chinese Mandarin, Tamil, Thai, Vietnamese, Russian and so far. As the participants are Chinese EFL learners, the Mandarin version was used in this study. Because some Chinese translation of items' options were not that exact, very small modification were made by the writer to make sure a better understanding for participants and more smooth evaluation. According to the suggestions mentioned in previous studies, that knowledge of 8,000 to 9,000 words may account for 98% text coverage and thus lead to independent reading (Dang & Webb, 2014; Laufer & Ravenhorst-Kalovski, 2010; Nation, 2006; Schmitt, Jiang, & Grabe, 2011; Laufer & Levitzky, 2017), of the original 140 VST items, only the first 90 were used in this study (see Appendix C for a full version). Hence, as it is considered that 10 items in each frequency band are assumed to represent 1,000 word families, then 90 items possibly represent 9,000 word families. The 90 items were presented randomly for participants. Even though the VST, as multiple choice test with four options, allows for a 25% chance of the test takers guessing the correct option regardless of whether the word's meaning is known or not, this study didn't use a fifth choice as include a fifth "I don't know" as suggested by some researchers (Bennett & Stoeckel, 2012; McDonald & Asaba, 2015; Zhang, 2013), as some other researchers found that even learners with similar vocabulary size used "I don't know" option differentially (e.g. Stoeckel, Bennett, & McLean, 2016). This study also didn't adopt the idea of Laufer and Tamiaviad-Levitzky (2017) in which participants were instructed not to answer items they did not know to minimize the effect of guessing, because it seems not that different from "I don't know" option. While it seems impossible to guess all successfully especially for some low-frequency words (McLean, Kramer, & Stewart, 2015; Stewart, 2014) and informed guessing may sometimes reflect partial knowledge (Nation & Webb, 2011), this study uses VST in an original four-option multiple choices state.

As for the L2 word association task, the stimulus words were selected from the VST with considering of a good balance of word frequency and part of speech. One noun, one verb and one

adjective from each frequency level were considered to be chosen, however, as there were no adjectives in k5 and k7 in the VST and no verbs in k9, the 24 stimulus words includes nine nouns, eight verbs, and seven adjectives, lastly. Each stimulus word with frequency level and word's part of speech is showed in Table 13. V means Verb. N means Noun. A means Adjective.

Table 13

Frequency Level and Word's Part of Speech of L2 Stimulus Words

frequency level	word (part of speech)
k1	see (V)
k1	period (N)
k1	poor (A)
k2	maintain (V)
k2	upset (A)
k2	patience (N)
k3	restore (V)
k3	jug (N)
k3	lonesome (A)
k4	candid (A)
k4	quiz (N)
k4	allege (V)
k5	weep (V)
k5	fracture (N)
k6	devious (A)
k6	accessory (N)
k6	veer (V)
k7	shudder (V)
k7	bloc (N)
k8	eclipse (N)
k8	authentic (A)

k8	mumble (V)
k9	perturbed (A)
k9	fen (N)

The depth of all stimulus words was evaluated by the Vocabulary Knowledge Scale (VKS) (Paribakht & Wesche, 1997). The VKS was employed to measure the participants' lexical knowledge for an individual word. This is a self-report rating scale and can measure initial vocabulary development according to performance on lexical items. For each word, the test produces five rating level indicating progressive degrees of lexical knowledge: (1) word unknown, (2) recognition of the word form, (3) vague word meaning, (4) word meaning, and (5) productive use (the highest degree among five). Despite the criticism that it might not reveal the knowledge differences only in the five scales (e.g., Wolter, 2001), it has been widely adopted in studies to measure L2 learners' initial vocabulary development for convenience of implementing. In addition, Read (2000) pointed out that it is sensitive and able to capture increases in vocabulary knowledge. Look at Table 14 for details of the five scales.

Table 14

Word Knowledge Assessed in VKS

Rating	Sample question	Word knowledge type
1.	I don't remember having seen this word before	Nil knowledge (word unknown)
2.	I have seen this word before, but I have little idea about its meaning.	Recognition of word form (partial knowledge)
3.	I have seen this word before, and I know it means _____. (synonym or Chinese translation)	Vague word meaning
4.	I know this word well. It means _____. (synonym or Chinese translation)	Word meaning
5.	I can use this word in a sentence: _____.	Productive use

4.3.3 Procedure

All participants were asked to complete the tasks online by accessing the URL of an online test service homepage “Wenjuanxing” where all experiment contents were put in by the author. Participants were asked to read a written consent with the aim of the experiment and the private and confidential clause on it, and sign it (enter the name) if he or she agrees to take part in the experiment and agrees the collocated data to be used in this study (see Appendix D). All of the participants sign the consent.

To start off, they were told that the aim of the experiment was to investigate the state of lexical network in L2 mental lexicon and to measure the vocabulary knowledge they have. Then the experiment began with written instructions in L1, followed by the tasks. The order of tasks was same for all participants: the L2 word association task, the VKS, and the VST were implemented in sequence; then a questionnaire was given to get information about their English learning background (see Appendix A, English Learning Background Questionnaire).

As for the L2 word association task, the stimulus words were ordered randomly. Prior to the actual L2 word association tasks, an example was given to familiarize the participants with the way of association. The examples are as follows: for the stimulus word “apple”, association words can be like App, red, fruit, orange, tree, sweet, eat, pie... A set of task materials with stimulus words were given and the participants’ task was to write down as many associated words as possible for each stimulus word within two minutes. Participants are noticed not to associate words in succession, in other words, not to take the words associated as a new stimulus word and associated from it. 15 blanks were prepared, among which five blanks were set as required and others were optional. If participants can only provide association words less than five, “0” is asked to be entered in to ensure the smooth running of the experiment. If there are more than 15 association words, the words more than 15 are asked to be put into the fifteenth blanks. Although nobody provided more than 15 words in this experiment actually.

Then in the VKS part, all stimulus word were presented whit five choices. Participants were requested to choose the best proper choice that accords with their own consideration about each words honestly. If the choice was 3, 4 or 5, the meanings or sentences in which the target words are used productively were requested to be answered, too.

In the VST, participants were required to choose the right L1 Chinese meaning from the four choices of each English word presented in a sentence. There was no time limitation to take the test.

Considering that it takes too long time to finish all contents once for some of them, participants were allowed to take a brief break between different parts of tasks. Though there are no clear recommendations regarding the length of time for taking such tests, most of the participants in our research completed all tests within about half and an hour to two hours except break time.

4.3.4 Statistical Analysis

The vocabulary size scores measured by the VST, and the depth of vocabulary knowledge measured by VKS were calculated for each participant. The VST scores were scored dichotomously, each correct item received 1 point, while wrong items received no point, so the scores ranged from 0 to 90. Even though the VKS is a kind of "combination of self-reported and demonstrated knowledge" (Wesche & Paribakht, 1996, p. 30) and which provides the depth of individual word knowledge scores on a scale of 1 to 5. Participants' self-reported scores were checked, and reevaluated by the writer according to their written contents for scale 3, 4 and 5. If the meanings which participants written were judged as wrong, the score will be modified from 3, 4 or 5 to 2. If the meaning is right but the sentence provided is wrong, the score will be modified from 5 to 4. According to Wolter (2001), of the 5 score scale, "a score of 5 can be thought to represent a word in the core of the mental lexicon, a score of 4 would represent a word that is fairly well known, a score of 3 indicates a word that is moderately well known, a score of 2 represents a word that is only slightly known, and a score of 1 indicates an unknown word, outside the mental lexicon." The VKS scores still ranged from 1 to 5 after modifying.

The category criterion in this study is the same as in Experiment 1 (see Table 5). Each of the associated word responses obtained were classified into seven categories of association according to the lexical semantic relationship between stimulus words and association responses: equivalent meaning, nonequivalent meaning, meaning and collocation, collocation, form and meaning, form and others. Response words were classified as others when there were responses which were just a copy of the stimulus words, or the responses are considered have no presumable semantic relation between the stimulus words and responses. Association English words, which are considered as spell miss and the real words of them are apparent, are corrected by the writer and counted as the word it should be.

As for the L2 word association task, all responses of the 24 items from all forty participants

were classified and labeled as Category 1 to 7 according to Table 5. A half of the data, that is data from 12 items, was drawn randomly from the 24 items, and the extracted data were divided in half (6 either) further randomly. All the extracted 12 items' data were evaluated by the author, the two sets of 6 items' data were evaluated by two other raters who are two participants of this experiment, respectively. In other words, both of them scored about a quarter of all the data. Considering that it is not that familiar for the two raters to evaluate such a kind of items, before the formal evaluation, an evaluation practice was carried out by making use of the data from two items of the other half items. The response items were evaluated by the author and other two raters together referring to the categorization criterion in Table 5. The online dictionaries (both English-English dictionary and English-Chinese dictionary) and online corpus SKELL which can check the words' examples in sentences, word sketch and similar words were used as some kinds of assistance. The raters' consistency was 82% between the author and rater A and was 86% between the author and rater B. The disputed cases for the data were discussed until agreement was reached. Because the inter-rater reliabilities were high enough according to Shweta, Bajpai and Chaturvedi (2015), the other half of the data (12 items) were evaluated by the author alone. Lastly, there was a total number of 4132 association responses for 24 stimulus word items from forty participants.

As for the grouping situation of HIGH or LOW, according to Nation's (2006) estimates, and as supported in several subsequent studies, 6,000 to 9,000 word families can account for 98% of text coverage, which, in turn, enables independent reading, the VST score 60 was determined to be used as a criterion of either it is HIGH or LOW vocabulary size. In other words, participants with a HIGH vocabulary size are considered as an independent reader in L2 English, while participants with a LOW vocabulary size are considered do not have enough vocabulary to read L2 materials independently. According to the level of participants' vocabulary size evaluated by the VST, the forty participants were divided into two groups with HIGH or LOW vocabulary size. There were 18 participants whose VST scores were below 60, and 22 participants whose VST scores were equal or higher than 60.

4.4 Results and Discussions

4.4.1 Vocabulary Size and Word Association

4.4.1.1 Correlation of Association Number of Each Category in L2 and the Vocabulary Size

Research Question 1: Does the vocabulary size (the breadth of vocabulary knowledge) affect Chinese EFL learners' word association responses?

The vocabulary size and the number of association responses in L2 were submitted to a correlation analysis. There was a significant but weak positive correlation of the vocabulary size and the number of association responses in L2 (r= .393, t (38) = 2.638, p= .012 < .05). The number of responses per participant in each category was counted then and the number of association responses of each category in L2 and the vocabulary size were submitted to another correlation analysis. In the result, there was a significant and moderate positive correlation between the number of association responses in L2 and the vocabulary size in Category 1, Category 2 and Category 3; a significant moderate negative correlation between the number of association responses in L2 and the vocabulary size in Category 6; a weak but significant positive correlation between the number of association responses in L2 and the vocabulary size in Category 4; and no correlation between the number of association responses in L2 and the vocabulary size in Category 5 and Category 7. More details of correlation analysis's results of association responses' number of each category in L2 and the vocabulary size are showed in Table 15 below.

Table 15

Correlation Analysis's Results of Association Responses' Number of Each Category in L2 and the Vocabulary Size

Category	r	t (38)	*p*
Category 1	.638**	5.103	.000
Category 2	.685**	5.795	.000
Category 3	.472**	3.298	.002
Category 4	.346*	2.274	.029
Category 5	-.117	-0.725	.473
Category 6	-.514**	-3.694	.001
Category 7	-.064	-0.394	.696

Note: ** means a moderate correlation and * means a weak correlation.

In short, there was a positive correlation of the vocabulary size and the number of association responses in L2 generally, and in Category 1, Category 2, Category 3, and Category 4, but a negative correlation in Category 6.

This means that generally speaking, as the participants' vocabulary size enlarging, the number of associations increases overall, the meaning-related links (equivalent meaning, nonequivalent meaning, meaning and collocation) increase, while the form-based links decrease.

4.4.1.2 Comparing the LOW and HIGH Vocabulary Size Groups: Number of Association Responses per Category

Research Question 2: Are there significant differences of the number and the type of association responses in L2 for participants groups of different vocabulary size (HIGH and LOW)?

The Welch's t-tests were conducted to determine if there was a statistically significant difference between the number of word associations of HIGH vocabulary size group (N=18) and that of LOW vocabulary size group (N=22). A *p* value less than 0.05 was considered statistically significant. Results revealed that the difference between the total number of association words of the two groups was not significant, t (27.643) =1.777, $p = .086 > .05$. The correlation coefficient was low, $r = .288$, and Cohen's d was moderate, d= 0.578. Considering the influence of Category 7 Others which was not that relevant, the number of association words except Category 7 of the two groups was conducted to the Welch's t-test again. Results revealed that the difference between the number of association words except Category 7 of the two groups was significant, t (32.277) =2.800, $p < .05$. The correlation coefficient was moderate, $r = .421$, and Cohen's d was high, d= 0.892. The number of association words of all seven categories of the two groups was conducted to the Welch's t-test and the result are showed in Table 16.

Table 16

The Difference of Association Responses' Number of Each Category in L2 Between Two Vocabulary Size Groups

	difference	standard error	t	df	*p*	correlation coefficient *r*	Cohen's d	Prep
Category 1	-20.172	4.666	-4.323	36.185	0.000	-0.558	-1.291	0.997
Category 2	-18.899	4.412	-4.284	37.332	0.000	-0.568	-1.325	0.998
Category 3	-2.298	0.836	-2.747	35.203	0.009	-0.410	-0.863	0.971
Category 4	-3.162	1.155	-2.736	30.230	0.010	-0.383	-0.796	0.961
Category 5	0.551	1.070	0.515	23.735	0.612	0.089	0.171	0.650
Category 6	19.657	6.831	2.878	25.259	0.008	0.442	0.947	0.981
Category 7	5.293	4.702	1.126	31.142	0.269	0.185	0.361	0.791

In a word, the results revealed that the difference between the total number of association words (except Category 7) of the two groups was significant, and more specifically, the difference between the number of association words of Category 1, Category 2, Category 3, Category 4, and Category 6 of the two groups was significant. For Category 1, Category 2, Category 3 and Category 4, the HIGH vocabulary group produced significantly more, while for Category 6 the LOW vocabulary size group produced significantly more. This result is in accord with the result of Research Question 1.

It is accessible that there will be a meaning-based to form-based shift as the vocabulary size enlarges. What worth noticing and should think about is the Category 4 "collocation only". The results showed a weak correlation between vocabulary size and associations linked with stimulus words by Category 4 "collocation only". This means the collocational links' increase may be not that obvious as the vocabulary size enlarges as meaning-based links are. Therefore, more attentions should be paid to collocation in foreign language learning. In fact, unlike equivalent meaning links which less likely to change from language to language as the universal conceptual background, as a kind of intralexical link that language users need to develop for an efficient lexical network, collocations, on the other hand, have been shown to vary considerably from language to language. Furthermore, this variation sometimes involves a measure of arbitrariness as there is often no logical or semantic reasons for why different languages use certain collocation expressions. Wolter and Gyllstad (2011) mentioned that some studies of learners' corpora demonstrated that even advanced language learners sometimes struggle to product appropriate

collocations in L2 use (for example, Bahns, & Eldaw, 1993). Unfortunately, however, there is a lack of the psycholinguistic research that has explored the psychological reality and the development of collocations in the L2 even though they may play an important role in structuring mental lexicon. The result of this study also showed that vocabulary size enlargement may facilitate the collocation learning, but not that directly and definitely, so much more attentions should pay to explore how the collocation in L2 are learned by foreign language learners and what can the teachers do to facilitate the development of collocational links in learners' mental lexicon constructing process, in addition to simple individual word learning in vocabulary acquisition research.

4.4.2 The VKS and Word Association

Research Question 3: Does the depth vocabulary knowledge affect Chinese EFL learners' word association responses? Is the number of association responses in L2 increases as the depth of knowledge of stimulus words increasing? Do response types change as well?

Considering the effects that depth of individual word knowledge may have on patterns of response type, the response patterns were analyzed within the categories designated by the VKS scores. The mean proportion of each response type within each VKS scale is shown in Table 17, *Figure*s 10. The total number of responses and numbers per category for each VKS scale is also given.

Table 17

Distribution of Association Category Based on VKS Level

	Category 1		Category 2		Category 3		Category 4		Category 5		Category 6		Category 7		Total	
VKS1	125	(19%)	153	(23%)	29	(4%)	8	(1%)	21	(3%)	172	(26%)	160	(24%)	668	100%
VKS2	159	(13%)	161	(13%)	26	(2%)	14	(1%)	56	(5%)	485	(39%)	327	(27%)	1228	100%
VKS3	84	(25%)	109	(33%)	18	(5%)	10	(3%)	30	(9%)	41	(12%)	40	(12%)	332	100%
VKS4	258	(33%)	213	(27%)	39	(5%)	34	(4%)	53	(7%)	79	(10%)	111	(14%)	787	100%
VKS5	400	(36%)	322	(29%)	23	(2%)	88	(8%)	39	(3%)	95	(9%)	150	(13%)	1117	100%

Table 17 shows the distribution of word association types based on the scale of vocabulary

knowledge as measured by VKS. On one hand, it demonstrates that responses to unfamiliar stimulus words (VKS1 nil knowledge and VKS2 word form) are non-relevant (Category 7) and form-based (Category 6) predominantly: Category 6 and Category 7, 332 (50%) out of 668 for nil knowledge and 82 (66%) out of 1228 for word form, respectively. Then, non-relevant (Category 7) and form-based (Category 6) responses decrease as the degree of knowledge of the stimulus words increases: VKS 3 vague word meaning (24%), VKS 4 word meaning (24%) and VKS 5 productive use (22%). The similar results among VKS 3, 4, and 5, on the other hand, indicate that after a certain stage as the words are no longer unfamiliar, even the degree of knowledge of the stimulus words increases, the association patterns do not change too much. On the other hand, when knowledge of the stimulus word become deep (VKS 4 word meaning or VKS 5 productive use), there are more meaning-based (for example, the five meaning-based categories: Category 1 to Category 5 occupy 78% for VKS 5 productive use words). These kinds of difference are showed in *Figure* 10 more directly and clearly.

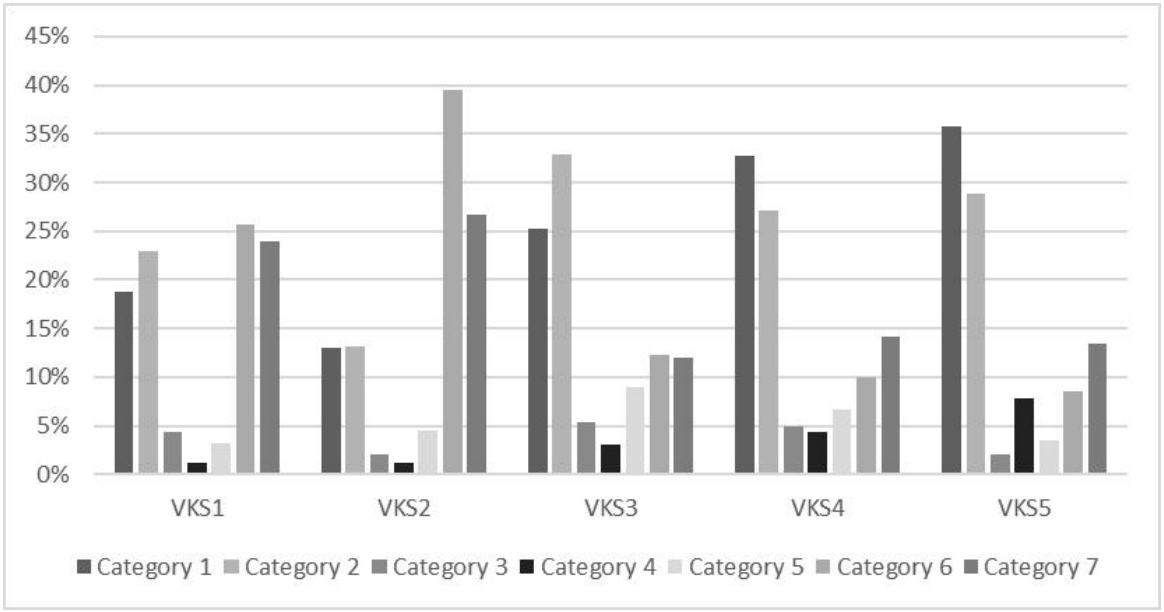


Figure 10.

Distribution of association category by VKS level

The results also show that form-based associations (Category 6) persist to some extent even for well-known words, which is in line with the finding of some previous studies (Li & Wang, 2016; Ma & Lee, 2019).

In comparing the distribution of responses within each VKS scale, the ANOVA with VKS scale as a between-subject factor was also conducted to find out whether there is any significant difference in the word association types for each level of VKS. The result showed that the main effect of VKS to Category 1 was significant, F (4,955) =44.659, p< .001, η_p^2=.158 R^2= .158, Adjust R^2= .154. As showed in Table 18, the main effect of VKS to other six association categories were also significant (p < .05).

Table 18

The Main Effect of VKS to Association Category

	F (4,955)	p	η_p^2	R^2	Adjust R^2
Category 1	44.659	.000	.158	.158	.154
Category 2	24.115	.000	.092	.303	.092
Category 3	2.495	.041	.010	.102	.010
Category 4	19.159	.000	.074	.273	.074
Category 5	5.133	.000	.021	.145	.021
Category 6	33.939	.000	.124	.353	.124
Category 7	5.567	.000	.023	.151	.023

In brief, the vocabulary knowledge of the stimulus word appears to be an important factor that influences the word association types. The meaning-based association increases, while form-based association decreases along with the word become better-known. It appears that although the form-based connections between words in the L2 mental lexicon do sometimes seem to take precedence over meaning-based connections for words that are not well known enough, this tendency changes as better understanding of individual words is gained, at which point the semantic connections become dominant. This does not mean, however, that formed-based connections are not important anymore in the L2 mental lexicon as we can see form-based associations (Category 6) remain to some extent even for well-known words at the level of VKS3, VKS4 and VKS5.

4.4.3 Comparing the Nouns, Verbs and Adjectives: The Proportion of Association Responses

per Category

Research Question 4: Does the stimulus words, especially their part of speech (nouns, verbs, or adjectives) affect the response types?

Firstly, numbers of association responses per word and per category and the average number of association responses for each word produced per participant are showed in Table 19.

Table 19

Number of Association Responses per Word and per Category in Experiment 2

frequency level	word (part of speech)	C 1	C 2	C 3	C 4	C 5	C 6	C 7	Total	Average number per participant
k1	see (V)	100	30	7	19	0	37	19	212	5
k1	period (N)	71	28	14	25	15	11	37	201	5
k1	poor (A)	58	75	1	13	6	19	18	190	5
k2	maintain (V)	77	5	0	28	5	24	47	186	5
k2	upset (A)	104	52	0	0	0	26	15	197	5
k2	patience (N)	23	58	0	0	20	21	58	180	5
k3	restore (V)	3	29	4	6	38	55	40	175	4
k3	jug (N)	43	35	39	3	0	68	18	206	5
k3	lonesome (A)	2	65	1	0	55	28	34	185	5
k4	candid (A)	47	24	0	3	1	47	42	164	4
k4	quiz (N)	71	39	9	26	2	24	20	191	5
k4	allege (V)	64	30	0	0	1	29	33	157	4
k5	weep (V)	26	89	0	2	6	40	33	196	5
k5	fracture (N)	30	40	0	1	23	30	21	145	4
k6	devious (A)	31	23	0	4	3	38	52	151	4
k6	accessory (N)	52	14	2	11	1	45	41	166	4
k6	veer (V)	32	42	4	0	1	40	45	164	4
k7	shudder (V)	30	45	0	0	2	32	34	143	4
k7	bloc (N)	27	46	12	2	0	54	31	172	4
k8	eclipse (N)	3	46	41	1	0	18	37	146	4
k8	authentic (A)	61	21	1	10	5	31	33	162	4
k8	mumble (V)	20	37	0	0	7	46	36	146	4
k9	perturbed (A)	33	35	0	0	8	37	22	135	3
k9	fen (N)	18	50	0	0	0	72	22	162	4
		1026	958	135	154	199	872	788	4132	

Note: C= Category

As the number of stimulus words of different part of speech is a little bit different originally, rather than the number, the percentage of different types' responses for stimulus nouns, stimulus verbs and stimulus adjectives are showed in Table 20 below.

Table 20

The Percentage of Different Types' Responses for Stimulus Words of Different Word's Part of Speech

	noun	adjective	verb
Category 1	22%	28%	26%
Category 2	23%	25%	22%
Category 3	7%	0%	1%
Category 4	4%	3%	4%
Category 5	4%	7%	4%
Category 6	22%	19%	22%
Category 7	18%	18%	21%
Total	100%	100%	100%

It shows that words in different part of speech may favor similar types of associations. The most frequent association responses belong to Category 1, Category 2, and Category 6 for nouns, verbs, and adjectives.

Figure 11 shows the distribution patterns for each part of speech more clearly. The results in both Table 20 and *Figure* 11 suggest that stimulus words' part of speech did not play an important role in determining the type of word association responses in this study.

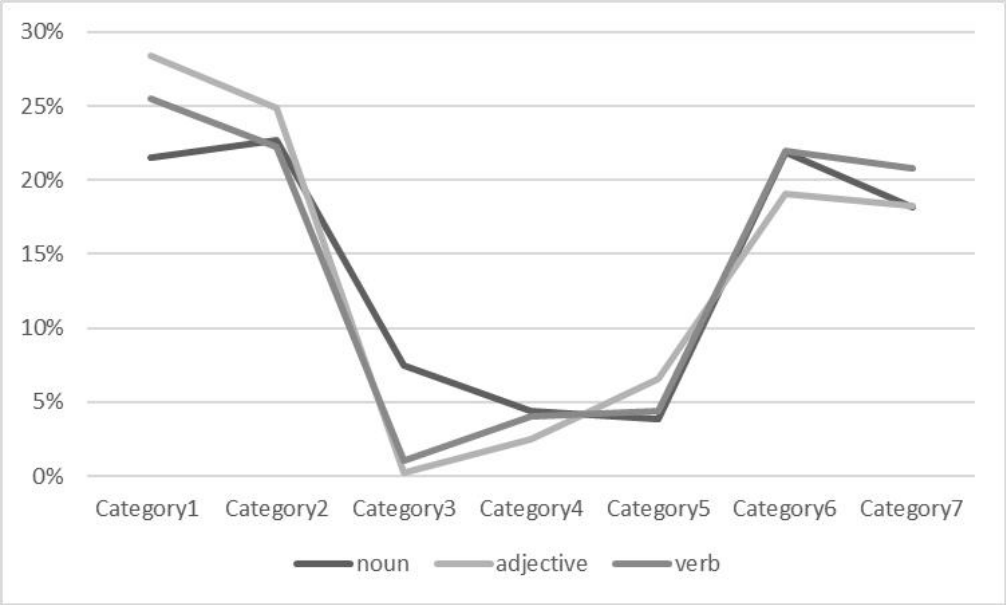


Figure 11.

The distribution patterns of association category for each part of speech.

This result is different from Experiment 1 and other researchers in previous study (e.g., Cronin, 2002; Nissen & Henriksen, 2006; Li & Wang, 2016; Zareva, 2011). The stimulus words' part of speech (verbs, adjectives, or nouns) actually showed different response results in these studies, in which nouns tend to elicit more paradigmatic responses while verbs more syntagmatic responses, and adjectives elicit both. In Experiment 1, stimulus verbs prompt a larger portion of responses in meaning and collocation (syntagmatic) category than stimulus nouns, while nouns prompt a larger portion of responses in equivalent meaning and nonequivalent meaning (paradigmatic) than stimulus verbs. Nissen and Henriksen (2006) showed that even when the stimulus words are well known to participants, nouns elicit more paradigmatic responses, while verbs and adjectives elicit more syntagmatic responses. Li and Wang (2016) similarly found that both verbs and adjectives lead to more syntagmatic responses than nouns. Zareva's (2011) result was slightly different from the two: nouns and adjectives product more paradigmatic responses, whereas verbs tend to product more syntagmatic responses.

The reason could be considered as follows:

One possible reason is that among the eight verbs used in this study (see 看, maintain 维持, restore 复原, allege 辩解, weep 哭, veer 改变方向或路线, shudder 发抖, mumble 咕哝) , most are intransitive verbs and can't take an objective directly in L2 or its homologous L1 translation

word even can't take an objective in L1 Chinese like 哭，辩解，发抖. Therefore, the words linked with them syntagmaticly is little originally. The second possibility is that generally speaking, the acquisition of collocations (syntagmatic responses) is difficult for L2 learners than the acquisition of single words. Hence, for L2 learners, to recall the synonyms, near-synonyms, superordinate, or hyponym for a word which is isolated is much easier than to recall a collocational linked word which is bound together with others, no matter what the part of speech is. Especially for some words that are not acquired completely yet for the learners, only the corresponding meaning of L1 is known, so getting through the meaning some words like synonyms are easier to be retrieved from the lexical network in the mind. However, further research is necessary to verify these postulates.

4.5 General Discussion

The relations among lexical items are highly complex and dynamic and the structure of the mental lexicon is further far from known-well. Among various kinds of links in the mental lexicon, collocational links seem especially difficult for L2 learners from the results of this study. Among many types of responses that have been investigated on word association tasks as a useful way of exploring the L2 lexicon (e.g., Meara, 1982; Soderman, 1993; Wolter, 2001, 2002; Fitzpatrick, 2006, 2007; Zareva, 2007), collocational responses comprise only one of them and are not paid enough special attention yet. Most of the psycholinguistic research to date on bilingual memory has been focused almost exclusively on interlexical or interlingual links between L1 and L2, with very little attention given to intralexical links in the L2 (French & Jacquet, 2004). Nonetheless, the little research about priming that has been done on collocational links indicates that they may play a vital role in structuring the mental lexicon in both L1 and L2. For example, Frenck-Mestre and Prince (1997) found a limited amount of priming for L2 collocational items, but as the collocational links in the study were combined with other types of syntagmatic links, it is impossible to know whether it was a significant priming for collocational links alone. Given the lack of research, and the fact that collocational connections seem particularly troublesome for L2 learners, to better understand the process of L2 collocations' development and the potential influence of L1 knowledge on L2 collocations at a psycholinguistic level is important.

4.6 Conclusion

4.6.1 Summary of the Findings

The present experiment continued the line of attempting to get a better understanding of what the L2 mental lexicon looks like in Experiment 1, with the underlying intention of trying to demonstrate that the L2 mental lexicon is not structured loosely or randomly but deeply related with both vocabulary size and the depth of vocabulary knowledge of each word. There is a correlation between the number of valid responses and vocabulary size and the difference between HIGH and LOW vocabulary size groups is significant. When the participants' vocabulary size enlarging, the number of associations increases overall, the meaning-related links (equivalent meaning, nonequivalent meaning, meaning and collocation) increase as well, while the form-based links decrease. According to the extent of how well a word is known, the types of association response changes. The meaning-based association increases, while form-based association decreases when the word become better-known. However, even there is a fair amount of support for the notion that while both the breadth and the depth of word knowledge of the stimulus word appear to be important factors that influence the type of word association, the part of speech which said to affect the associations of equally well-known words, the part of speech effect were not work based on the results in this experiment. The possible reasons for the result were discussed. The collocational links in L2 mental lexicon was focused.

4.6.2 Pedagogical Implications

First of all, although vocabulary teaching is still usually neglected by many teachers in teaching practice nowadays, as the results showed that the vocabulary knowledge plays a really important role in lexical network, vocabulary teaching should be much accounted of in L2 teaching.

Although the teaching methods is beyond the scope of this study, there are some points to indicate the rough direction of teaching. One point is to take advantage of the system of word association and to teach words linked in the network by various kind of connection categories at the same time. For example, teachers can pay more attention to the links appeared easily in word association of L2 learners or links that is important (for example, appearing frequently in L1 association) but rarely appeared in L2 learners' association, and think out a proper way to teach

them differently. The other point is about the difficulty of collocation acquisitions revealed in this experiment. Teachers may need to pay more attention to collocation teaching in addition to usual vocabulary teaching. When teaching a new word, the teachers may tell students which words (both learned words and new words are possible, although too much new words may lead to weariness of learning, which should notice in practice) collocate with the word. This may help the L2 learners form links between the new word and other known words in their mental lexicon.

Another point worth noting is that this study shows that L2 learners tend to produce form-based word associations on encountering unfamiliar words and some of them may do not have sufficient vocabulary to deal with L2 reading. It has been shown that vocabulary knowledge is highly correlated with reading in one's L1 and even more so with reading in L2. Some researchers even argued that learning to read in a L2 centrally involves learning words. It is essential to help L2 learners of English enlarge their vocabulary size at least to the necessary extent in order to achieve a threshold level of lexical coverage, which is the percentage of word tokens they understand in a given L2 reading text. To some extent, the development of L2 vocabulary knowledge and L2 reading competence appears to in a mutual promotion relationship. Effective L2 reading helps develop vocabulary knowledge and enough vocabulary knowledge ensures a smooth L2 reading process and furthermore facilitates the L2 reading process. In teaching practice, this kind of information will be helpful for foreign language teachers to make an efficient and good balanced teaching plan.

4.6.3 Limitations and Recommendations for Further Research

This study has some certain limitations that need to pay attention to in further research.

First, the raters of Experiment 2 were the writer and two participants. However, three raters are all EFL learners, but the judgment of whether a pair of stimulus words and association responses belongs to a specific category, collocation category, for example, in a foreign language is difficult for a L2 learners sometimes, and may be different from native raters. Although it is difficult to say which one is better, as the native speakers may know the language better, while the L2 speakers know the participants better to some extent. An approach like the retrospective interview in Fitzpatrick (2006) which participants themselves coding their own responses according to a specific category criterion after word association tasks by themselves may be useful,

however, there may be some practical difficulties, for example, in this experiment, the burden may be too heavy for participants to evaluate themselves and may make matters worse as the participants gathering is very difficult already.

Second, learn a lesson from Experiment 1, to exclude the effect of dual part of speech of word like “break”, the part of speech was labeled in Experiment 2. However, as the prevalent polysemy in language, in addition to the common meaning, some participants know the rare meaning of some words and the association become different. It is a new problem we found this time. It is necessary to consider about to what extent the “free” should be in the free word association tasks to aim for more precise and reliable data in future studies.

CHAPTER 5 THE PRESENT STUDY: EXPERIMENT 3

Putting all results of Experiment 1 and 2 together, a collocation decision task was implemented in Experiment 3, to gain insight into the way of representation of L1 and L2 for Chinese English learners in mental lexicon in more depth, especially to know more about the strength and directions of links between words (both intralingual (collocational words in one language) and interlingual (equivalent words in two languages)) in the network depicted in Experiment 1. Then, the lexical decisions task with experimental priming paradigm under masked condition was implemented, to investigate how semantically related word pairs and translation word pairs would be recognized by the learners, to provide more evidence to disputes about either asymmetry or symmetry between L1 and L2, and if possible, to figure out the relationship among concepts, L1 representation, and L2 representation. Based on Experiment 1, collocations concerning “drink”, “open”, which in L1 Chinese have a broader semantic category and “wear”, “break”, which in L2 English have a broader semantic category were used in collocation decision task, while more words which have either kind of relations between L1 and L2 semantic category (L1>L2 or L1<L2) were selected as items for lexical decisions task.

5.1 Research Questions

For the purposes stated at the beginning of this chapter, the research questions of present study were posited as follows:

Research Question 1: Does the semantic category difference between learners’ L1 and L2 influence their L2 English vocabulary learning, same as Experiment 1 and previous studies?

Research Question 2: Is the strength of links between words different according to the Reaction Time (RT) data? Can it tell more details on the basis of the network figured in Experiment 1?

These two research questions above are going to be researched by a collocation decision task.

Research Question 3: Is the priming effect from L2 to L1 direction and L1 to L2 direction different? Is the priming effect different according to different relationships between stimulus words and target words (translation, related, unrelated, and translation pairs with non-word targets)? How does it differ?

Research Question 4: Are there any differences of translation priming between different semantic category?

Research Question 5: Are there any differences in translation priming effect between participants of HIGH and LOW proficiency?

These three questions are studied by a lexical decisions task with experimental priming paradigm under masked condition.

5.2 Hypotheses

Hypotheses are made according the results in literature and the features of experimental design in this study.

Hypothesis 1: Same as Experiment 1 and previous studies, the semantic category difference between learners' L1 and L2 are expected to influence their L2 English vocabulary learning.

Hypothesis 2: The RT data will reflect the strength of links to some extent for a strong link leading to a faster response and vice versa.

Hypothesis 3: It is said that there are translation priming both from L1 to L2 and from L2 to L1, however, the effects from L1 to L2 were significantly stronger than the effects from L2 to L1 in most previous studies. Therefore, we expected a stronger priming from L1 to L2 than the reverse in this study, too. As for the relationship between stimulus and target words, the priming effect may be different according to different relationships between stimulus words and target words (translation, related, unrelated, and translation pairs with non-word targets).

Hypothesis 4: There may be some differences of translation priming between different semantic category.

Hypothesis 5: Some differences in translation priming effect between participants of HIGH and LOW proficiency are expected.

5.3 The Main Study

The present study is trying to investigate the way in which Chinese English learners represent their two languages in mental lexicon, especially the strength and directions of links between both intralingual and interlingual words, by using collocation decision task and cross-language priming lexical decision tasks (translation, semantic, unrelated, and non-word translation).

5.3.1 Participants

Twenty Chinese EFL learners (ten females, and ten males) took part in the study. Four of them are university students (three students are at the fourth year and one is at the third year) and others have already graduated from universities and entered graduate schools as a master or doctoral student or are at work. All participants with Chinese Mandarin as their mother tongue have learned English as a foreign language for about 14 years in average. Participants were an average of 26 years old (range: 20—31).

5.3.2 Materials

5.3.2.1 Collocation Decision Task

In this task, both L1 Chinese and L2 English collocations are included. In the English part, collocations of words "drink" and "open" which were considered as representatives of English semantic category narrower than Chinese "喝 he" and "打开 da kai"; and words "wear" and "break" which were considered as representatives of English semantic category larger than Chinese "穿 chuan" and "打破 da po / 弄坏 nong huai" are included. There were 40 items in all with 20 correct collocations and 20 incorrect collocations. In the Chinese parts, the translation equivalents of words in the English parts are used. There were 32 items in all with 20 correct collocations and 12 incorrect collocations (Referring Appendix E for a list of items).

5.3.2.2 Cross-language Priming Lexical Decision Task

In this task, both direction of L2 English - L1 Chinese and L1 Chinese - L2 English are included. Chinese verbs "喝 he", "打开 da kai", "借 jie", "介绍 jie shao", "看 kan", "笑 xiao", "听 ting"; nouns "帽子 mao zi", "桌子 zhuo zi"; and adjectives "高 gao", "大 da", "小 xiao", were considered as representatives of Chinese semantic category larger than English semantic category of equivalent words; and Chinese "穿 chuan" and "打破 da po / 弄坏 nong huai", were used as representatives of Chinese semantic category narrower than English ("wear" and "break").

In the L2 English - L1 Chinese part, there were translation pairs, semantically related pairs, unrelated pairs, and translation pairs with non-word targets as Chinese single characters or two-character phrases being targets to be recognized. All the Chinese words mentioned above

were used as target words. Their English equivalent words were stimulus words in translation pairs. The English equivalent words were produced by translate the Chinese items into English by the author. Other three Chinese English learners (the VST score was 73, 55, and 42, respectively) were asked to confirm whether they know the English items or not and to write the Chinese translations of them to make sure the pairs are really equivalent between two languages. For translation pairs with non-word targets, the Chinese non-word targets were created by using a Japanese Kanji which do not exist in Chinese when the translation equivalent is one character or combining two Chinese characters together to make a nonsense two-character Chinese phrase when the translation equivalent is a phrase with two characters (Referring Appendix F). Semantically related pairs', and unrelated pairs' lists were made by the author generating Chinese targets according to the English stimulus words. The three Chinese English learners were asked to checked if the word pairs were semantically related or not, items which all people gave the same answer were used, while items with different answers were discussed by the author and the rater, then changed and checked by other raters until to be used.

In the L1 Chinese - L2 English part, similarly, there were translation pairs, semantically related pairs, unrelated pairs, and translation pairs with non-word targets as English words being targets to be recognized. While translation pairs, semantically related pairs, and unrelated pairs were just the opposite version of L1 Chinese - L2 English part, non-word targets were created by changing one letter in a real English word this time. Thus, all English nonwords were pronounceable (Referring Appendix F for a list of items).

5.3.2.3 VST

Participant's vocabulary size (i.e., the breadth of word knowledge, how many words someone knows) was measured by the Vocabulary Size Test (VST) (Nation and Beglar, 2007) the same as the Experiment 2 (Referring **4.3.2** for the details).

5.3.3 Procedure

The experiment was designed by the Super Lab 6. Then an experiment Package made by Super Lab Remote was sent to participants online. Every participant participated the experiment individually. When the experiment was implemented, participants were connected with the author

online face to face or by voice at least in the process. To start off, the author talked the instructions and highlighted important notes orally. Then participants were instructed to the download and open the Package file made by Super Lab, drag the file to the desktop and run that Remote Package. While the task running, participants were asked to read carefully and follow the on-screen instructions. They were told to ask whatever they feel confused at any time online before the experiment or at the break time between trails. When the task was complete, the app automatically exits. At this point, participants were asked to send a file named “Task Data.txt” on their desktop to the author. Afterwards, participants were told that they can delete the Package folder and the file “Task Data.txt” from their computers.

In the collocation decision tasks, stimuli were presented on the screen by Super Lab Remote, which also recorded reaction times and error data automatically. Each part started with the presentation of a mark “+” on the center of the screen for 500ms, then followed by the target collocations for 300ms per item, there was nothing presented between one item and another. Participants were instructed to decide whether each item presented was a correct collocation in that language (Chinese or English), and respond by pressing either a “z” button (“No”) or a “m” button (“Yes”) on the keyboard by the finger of left and right hand, respectively. Word pairs were presented randomly. Each participant completed five practice trials before the formal experiment.

In the lexical decision tasks, stimuli were presented using Super Lab Remote, which recorded reaction times and error data, too. Each trial started with the presentation of a forward mask “#####” for 500ms, followed by the prime stimulus for 250ms, a backward mask “&&&&&” for 50ms, and then the target words for 3000ms at most. Participants were instructed to decide whether each item presented was a real Chinese character or a word (two-character string), or a real English word, then to respond by pressing either a “z” button (“No”) or a “m” button (“Yes”) on the keyboard by the finger of left and right hand respectively. Word pairs were presented randomly. Each participant completed ten practice trials before the formal experiment.

All participants were asked to complete the VST and a questionnaire before or after the online tasks. The questionnaire in order to get information about participants’ English learning background (see Appendix A, English Learning Background Questionnaire) was finished online by accessing the URL of an online test service homepage “Wenjuanxing” with all contents set by the author.

Participants were asked to read a written consent with the aim of the experiment and the private and confidential clause on it, and sign it (enter the name) if he or she agrees to take part in the experiment and agrees the collocated data to be used in this study (see Appendix D). All of the participants sign the consent.

5.4 Statistical Analysis and Results

5.4.1 Collocation Decision Task

CORRECT (C), ERROR (E) or No Response (NR) decisions and all Reaction Times (RTs) were recorded by the Super Lab automatically.

5.4.1.1 drink (L1>L2)

The decision and Accuracy / Error Rates (ERs) concerned "drink" of both languages are presented in Table 21 below.

Table 21

The Decision Results of C, E, NR, and Accuracy / Error Rates (ERs) Concerning "drink"

Correct or not				C		E		NR				C		E		NR	
Only L1	drink	soup	*	2	10%	17	85%	1	5%	喝 汤		19	95%	1	5%	0	0%
Only L2	eat	soup		8	40%	11	55%	1	5%	吃 汤	*	18	90%	2	10%	0	0%
Both Not	eat	water	*	15	75%	5	25%	0	0%	吃 水	*	18	90%	2	10%	0	0%
Both OK	drink	water		20	100%	0	0%	0	0%	喝 水		19	95%	0	0%	1	5%
Only L2	take	medicine		17	85%	3	15%	0	0%								
Only L1	drink	medicine	*	9	45%	9	45%	2	10%	喝 药		17	85%	3	15%	0	0%
Only L1	eat	medicine	*	3	15%	17	85%	0	0%	吃 药		17	85%	2	10%	1	5%

For "drink water" and "喝水", which are equivalent and are correct in both L1 and L2, almost all participants judged correctly, with a high accuracy of 100% and 95%, respectively. For "eat water*" and "吃水*", which are equivalent and both are not acceptable in L1 and L2,

participants judged also with a high accuracy of 75% and 90%, respectively.

It is clear that the accuracy of items of “drink soup*,” “drink medicine*,” and “eat medicine*,” which needed to be rejected and “eat soup,” which is actually the correct L2 expression is really low (lower than 45%, some even just 10%) comparing to other items. As for the high error rates of these L2 English items, the analysis of only correct data is nonsense. In other words, the ERROR data is really important in this situation.

The results of judgment of “C”, “E” or “NR” patterns and the RTs of all participants can be found in Appendix G G1. Table 22 below only shows the data of some representative participants.

Table 22

The “C”, “E” or “NR” patterns and the RTs of “drink” for some representative participants

	drink soup*	eat soup	喝 汤	吃 汤*	drink water	eat water*	喝 水	吃 水*	take medicine	drink medicine*	eat medicine*	喝 药	吃 药
2	E 1186	C 1008	C 850	C 760	C 1226	C 1209	C 999	C 974	C 1299	C 1610	E 1199	E 756	C 1208
3	C 2584	C 1855	C 1145	C 1145	C 1746	C 1254	C 926	E 1431	C 2025	NR 0	E 2183	C 1322	C 1456
4	E 1346	C 1579	C 751	C 1233	C 848	E 845	C 738	C 1559	C 963	E 1103	E 1254	C 825	C 2848
6	E 1092	E 713	C 849	E 548	C 65	C 2561	C 1556	C 736	C 1148	C 1689	E 1272	C 1144	C 968
9	E 1553	E 1688	C 1554	C 1009	C 1174	C 1365	C 623	C 880	C 1031	C 1616	E 1373	C 1454	C 2087
10	E 1314	E 1787	C 15	C 865	C 1409	C 1759	C 702	C 1052	C 1273	C 1531	C 1720	C 10	C 745
11	E 1505	E 1895	C 1151	C 1637	C 1140	C 1201	C 709	C 869	E 1617	E 1462	E 1701	C 965	C 1156

Overall, for “drink water” and “eat water”, beside five participants misjudged “eat water*”, all judgments were correct, while for the equivalent Chinese items “喝水” and “吃水*”, three participants misjudged or gave no response to either of the two items (See Appendix G G1). Average of all correct data were calculated and showed in Table 23 below.

Table 23

The Average RT of Items Concerning “drink water”

	drink water	eat water*	喝水	吃水*
Average RT	1160.9	1535.533	1075.474	1156.444

It takes more time to judge “eat water*” as WRONG than judge “drink water” as CORRECT. The average RTs of Chinese is faster than English items.

For “drink soup*” and “eat soup”, only one participant correctly judged both items, eleven participants misjudged both items, six participants misjudge “drink soup*” but judged “eat soup” correctly, and two judged one correctly but gave no response to the other (1 C NR / 1 NR C); while for the equivalent Chinese items “喝汤” and “吃汤*”, only three participants misjudged one of the two items. The average RT of items concerning “eat soup” are showed in Table 24 below according to the response patterns.

Table 24

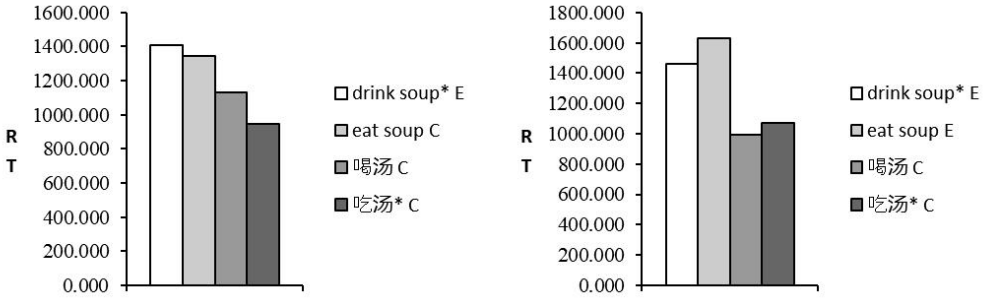
The Average RT of Items Concerning "eat soup"

	drink soup *	eat soup	喝 汤	吃 汤 *
3	C	C	C	C
RT	2584	1855	1145	1145
	E	C	C	C
2	1186	1008	850	760
4	1346	1579	751	1233
8	1248	1392	906	928
13	1526	1270	2090	885
16	1741	1473	1073	940
Average RT	1326.5	1312.25	1149.25	951.5
	E	E	C	C
5	1901	1408	954	890
7	1754	2597	539	1231
9	1553	1688	1554	1009
10	1314	1787	15	865
11	1505	1895	1151	1637
14	915	1048	635	650
15	1749	1317	2354	796
17	1383	1365	1004	1932
20	1090	1544	743	657
Average RT	1462.667	1627.667	994.3333	1074.111

For participant No.3 who correctly judged all items, the RTs of Chinese is faster than English items, and it takes more time to judge "drink soup*" as WRONG than judge "eat soup" as CORRECT.

The average RTs of participants who correctly judged both Chinese items and gave responses

to both English items but misjudge “drink soup*” while judged “eat soup” correctly (Figure 12, a) or misjudged both English items (Figure 12, b), respectively are compared in Figure 12. For both groups of participants, the average RTs of Chinese is faster than English items.



a *b*

Figure 12.

The average RT concerning “eat soup”.

According to the results above, the network of L1 and L2 taking “drink‧喝(water, soup) as examples are showed in *Figure* 13, 14, and 15. The intralinguistic links are showed by black solid lines, and with a RT (a shorter RT = a stronger solid line). The interlinguistic links are showed by dotted lines. Some inappropriate associations which negative evidence is considered necessary are showed by red lines. The rectangles or ellipses with the same color means equivalence of two words in two languages.

The network of L1 and L2 taking “drink‧喝(water, soup)” as examples for Participant No.3 who correctly judged all items concerning “eat soup” is showed in Figure 13.

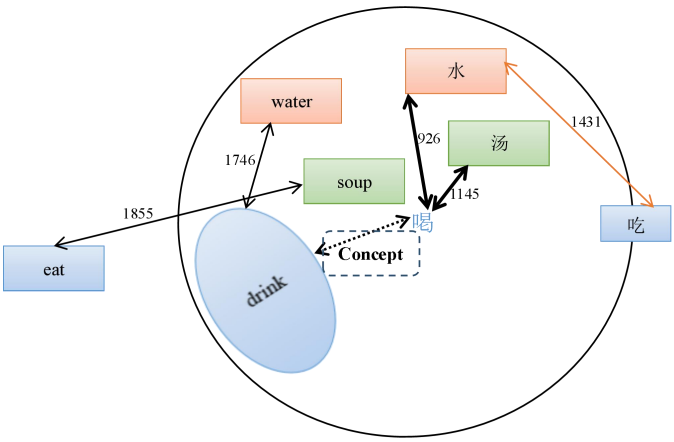


Figure 13.

The network of L1 and L2 taking "drink·喝(water, soup)" as examples for participant No.3.

For 5 participants who misjudge "drink soup*" but judged "eat soup" correctly, the average data was calculated, however, the original data showed that the RT patterns are diverse. This means that the average data is nonsense. Therefore, we take participant No. 2 who correctly judged all items concerning "drink water" and with the highest VST score in the group as an example. The results showed that it takes more time to judge "drink soup*" as CORRECT than judge "eat soup" as CORRECT, while it takes more time to judge "喝汤" as CORRECT than judge "吃汤*" as WRONG.

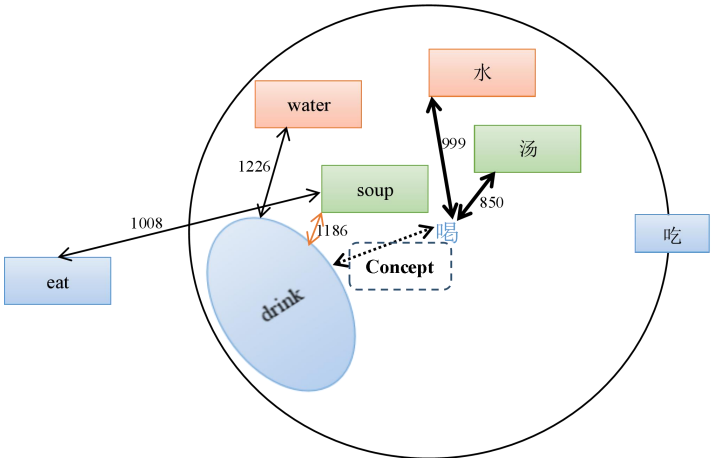


Figure 14.

The network of L1 and L2 taking "drink·喝(water, soup)" as examples for participant No. 2.

For nine participants who misjudged both items, the average data was calculated, however, the original data showed that the RT patterns are diverse, too. Therefore, we take Participant No. 9 who correctly judged all items concerned “drink water” and with the highest VST score in the group as an example. The results showed that it takes less time to judge “drink soup*” as CORRECT than judge “eat soup” as WRONG, while it takes more time to judge “喝汤” as CORRECT than judge “吃汤*” as WRONG. This result is interesting, because the participant takes more time to confirm correct items than to reject the wrong one in L1.

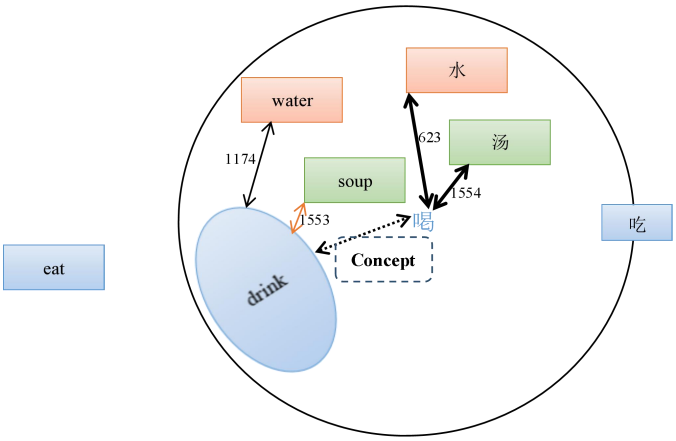


Figure 15.

The network of L1 and L2 taking “drink‧喝(water, soup)” as examples for participant No. 9.

For “take medicine”, “drink medicine*”, and “eat medicine*”, and the equivalent Chinese items “喝药” and “吃药”, six participants misjudged or gave no response to one of Chinese items, therefore, only other 15 participants’ data were considered. Among them, only two participants correctly judged both items, nine participants correctly judged “take medicine” but misjudged “drink medicine*” (4) or both (4) of another two English items or gave no response to “drink medicine*” (1), three participants misjudge all three English items.

The average RTs of participants who correctly judged both Chinese items and gave responses to all English items (correctly judged all items; misjudged all items; correctly judged “take medicine” but misjudged “eat medicine*” or both of another two items, respectively) are compared. For all participants, the RTs of Chinese seem faster than English items.

Table 25

The Average RT of Items Concerning "take medicine"

	take medicine	drink medicine *	eat medicine *	喝 药	吃 药
	C	C	C	C	C
10	1273	1531	1720	10	745
18	1799	1694	1551	853	714
Average RT	1536	1612.5	1635.5	431.5	729.5
	C	E	E	C	C
4	963	1103	1254	825	2848
5	1070	2936	2345	1579	730
7	1701	2396	1684	782	548
19	1909	1092	1260	1507	1082
Average RT	1410.75	1881.75	1635.75	1173.25	1302
	C	C	E	C	C
6	1148	1689	1272	1144	968
9	1031	1616	1373	1454	2087
12	2029	1229	1975	824	778
17	2657	1410	1617	1290	1180
Average RT	1716.25	1486	1559.25	1178	1253.25
	E	E	E	C	C
8	1417	1493	1027	704	1114
11	1617	1462	1701	965	1156
13	2497	2161	1261	1388	1608
Average RT	1843.667	1705.333	1329.667	1019	1292.667

Similar to items concerning "drink soup*" the average data was calculated, however, the original data showed that the RT patterns are diverse. Therefore, we take one participant who correctly judged all items concerned "drink water" and with the highest VST score in the group as

an example to show the network of L1 and L2 taking “drink·喝 (water, medicine) as examples.

Participant No. 10 as a represent of participates who correctly judged all items, it takes more time to judge “drink medicine*”, and “eat medicine*” as WRONG than judge “take medicine” as CORRECT.

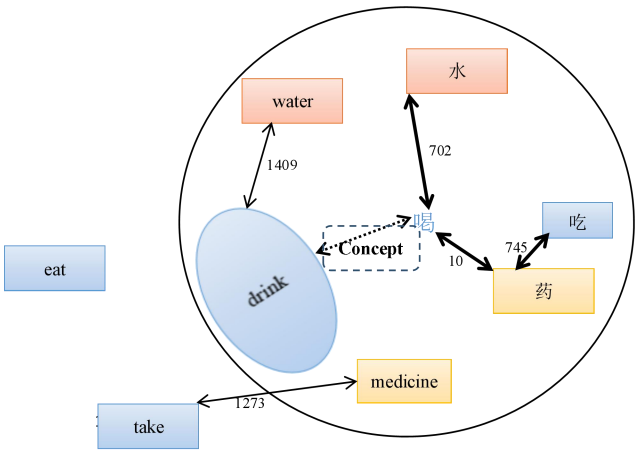


Figure 16.

The network of L1 and L2 taking “drink·喝(water, medicine)” as examples for participant No. 10.

Participant No. 11 as a represent of participates who misjudged all items, it takes more time to judge “take medicine” than “drink medicine*” and “eat medicine*”.

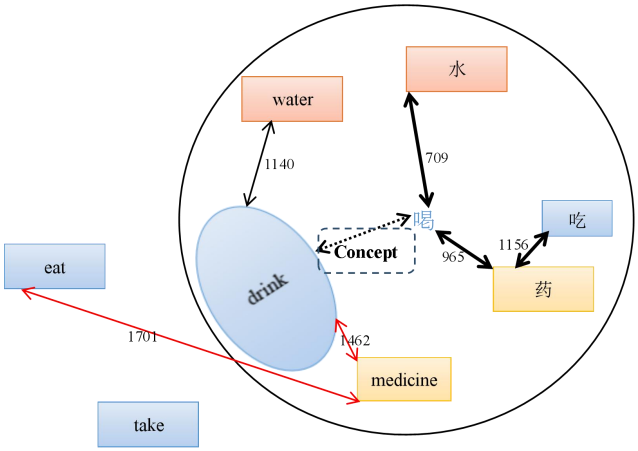


Figure 17.

The network of L1 and L2 taking “drink·喝(water, medicine)” as examples for participant No. 11.

Participant No. 6 as a represent of participates who correctly judged “take medicine” but misjudged “eat medicine*” only, it takes less time to judge “take medicine” as CORRECT than “drink medicine*” as WRONG and “eat medicine*” as CORRECT.

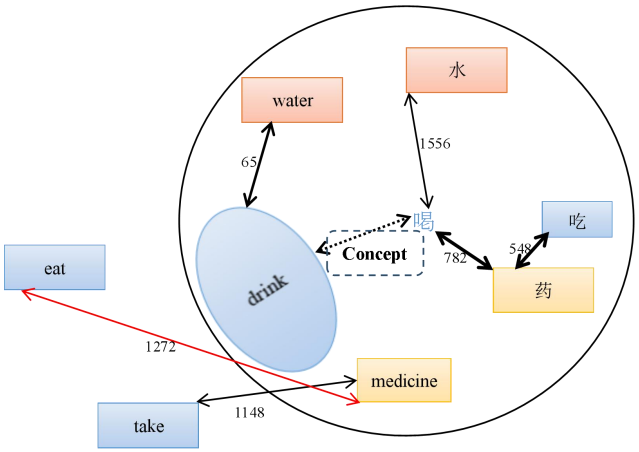


Figure 18.

The network of L1 and L2 taking “drink·喝(water, medicine)” as examples for participant No. 6.

For participates who correctly judged “take medicine” but misjudged both of another two items, no one correctly judged all items concerned “drink water”, therefore even misjudged “eat water*” as CORRECT, Participant No. 4 with the highest VST score in this group is selected as a represent. The results showed that it takes less time to judge “take medicine” as CORRECT than “drink medicine*” and “eat medicine*” as CORRECT.

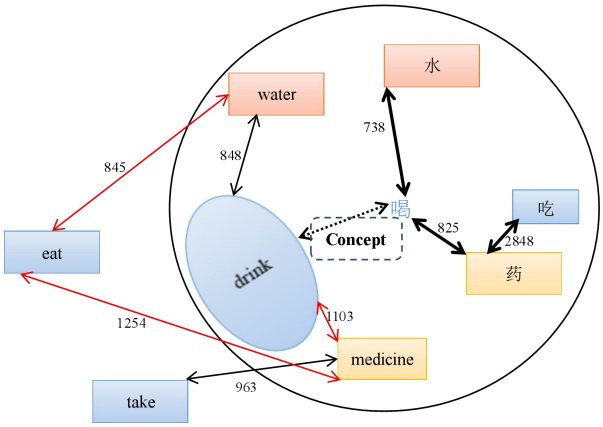


Figure 19.

The network of L1 and L2 taking “drink·喝(water, medicine)” as examples for participant No. 4.

In a word, the low accuracy data of some items makes the analysis of RT complicated, because the RT of CORRECR response and ERROR response should not be treated in the same way and is difficult to be compared directly. Then the average data are not representative in some cases, so the network of representation of participants for every typical situation were depicted. Results showed that although not only the low proficiency learners but also the high proficiency learners were still not able to reject the wrong expression completely; some learners have the correct L2 links in their mental lexicon while others do not.

5.4.1.2 open (L1>L2)

The decision and accuracy / error rates (ERs) concerned "open" of both languages are presented in Table 26 below.

Table 26

The Results of C, E, NR, and Accuracy of Items Concerning "open"

	C		E		NR	
open + door	19	95%	1	5%	0	0%
打开 + 门	20	100%	0	0%	0	0%
open + TV *	2	10%	18	90%	0	0%
turn on + TV	19	95%	1	5%	0	0%
打开 + 电视	19	95%	1	5%	0	0%
open + letter *	2	10%	17	85%	1	5%
unfold + letter	16	80%	4	20%	0	0%
打开 + 信件	19	95%	0	0%	1	5%

For "open + door" and "打开 + 门"，which are equivalent and both are correct in L1 and L2, almost all participants judged correctly, with a high accuracy of 95% and 100%, respectively. It is clear that the accuracy of items of "open TV*," and "open letter" which needed to be rejected is really low (10% for both items) comparing to other items.

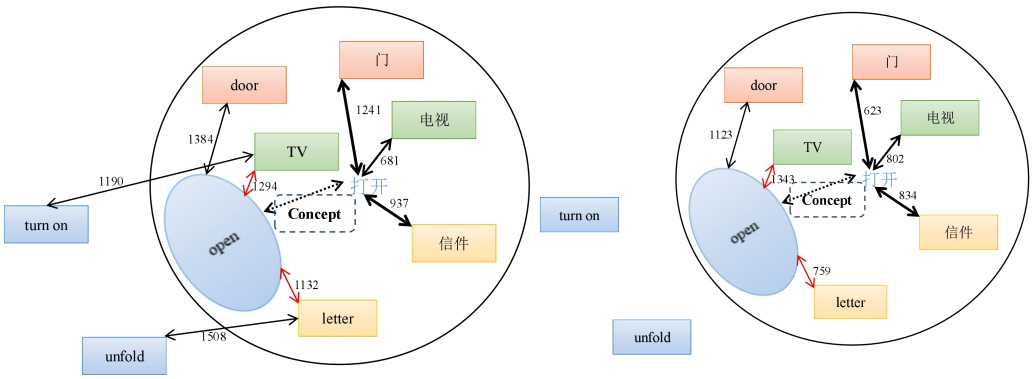
Overall, two participants misjudged or gave no response to either one of Chinese items, and only one participant misjudged “open door”, so other 17 participants’ data were considered. Among them, ten participants misjudged both “open TV*” and “open letter*” but judged all others correctly; three participants misjudged either “open TV*” or “open letter*” but judged all others correctly; four participants misjudged or gave no response to one or more items other than “open TV*” and “open letter*”. The results of judgment of “C”, “E” or “NR” patterns and the RTs of all participants showed in Appendix G G2. The average RTs were calculated for different patterns respectively and the data are showed in Table 27.

Table 27

The RTs of “open” in Different Decision Patterns

	open + door	打开 + 门	open + TV *	turn on + TV	打开 + 电视	open + letter *	unfold + letter	打开 + 信件
	C	C	E	C	C	E	C	C
2	969	704	832	1049	2031	982	1726	727
4	1384	1241	1294	1190	681	1132	1508	937
5	715	1316	880	354	1734	1300	1531	1041
9	932	1598	1005	1165	1506	1060	1499	997
10	906	714	966	1057	968	1333	2359	765
11	1012	825	1093	1248	722	1425	2192	774
12	1629	1378	1005	1211	845	1509	1274	1243
14	1009	795	790	801	726	1055	1185	729
15	1424	1410	769	1635	1554	1381	2730	1410
19	1019	819	797	1611	1715	1066	2115	1378
Average	1099.9	1080	943.1	1132.1	1248.2	1224.3	1811.9	1000.1
	C	C	C	C	C	E	C	C
1	1091	1261	288	1521	893	1805	2965	1393
16	1521	940	2018	1321	856	1709	2769	1652
Average	1306	1100.5	1153	1421	874.5	1757	2867	1522.5
	C	C	E	C	C	E	E	C
3	1395	1500	1406	1771	973	1885	2283	942
8	990	709	737	1358	649	1035	1474	776
Average	1192.5	1104.5	1071.5	1564.5	811	1460	1878.5	859
	C	C	E	C	C	C	C	C
6	1101	579	741	950	906	1312	1356	1856
	C	C	E	C	C	NR	E	C
17	1360	1382	1049	1223	1146	0	1742	1958
	C	C	E	E	C	E	E	C
20	1123	623	1343	2051	802	759	779	834

Similar to items cornering “drink”, although the average data was calculated, the original data showed that the RT patterns are diverse. Therefore, we take Participant No.4 with a relatively high VST score, who showed a faster RT to Chinese than to English in the same response pattern group and misjudged both “open TV*” and “open letter*” but judged all others correctly (Figure 20, a) and Participant No.20 with a relatively low VST score, who misjudged all “open TV*,” “turn on TV,” “open letter*,” and “unfold letter” (Figure 20, b) as examples to show the network of “open‧打开 (door, TV, letter) in L1 and L2 as an example.



a b

Figure 20.

The networks of L1 and L2 taking “open‧打开” as examples for participant No. 6 and No.20, respectively.

It is clear that even both of the two participants have wrong collocational links in their lexical network, there were links for correct collocations “turn on TV” and “unfold letter” in the network of Participant No.4 who have a larger vocabulary size but not in the network of Participant No.20 whose vocabulary size score is lower.

5.4.1.3 wear (L1<L2)

The accuracy of items concerning “wear” was showed in Table 28.

For all correct Chinese items, the accuracy was very high, while for wrong Chinese items, the accuracy was lower comparatively, but still higher than 55%. For all correct English items, even when the collocations were different between L1 and L2 and only is acceptable in L2 (items like

"wear glasses" and "wear necktie"), the accuracy was still high (90% and 70%, respectively). The accuracy of "wear earrings" (50%) was also not that low compared to items like "drink soup*" (10%) and "eat medicine*" (15%).

Table 28

The Results of C, E, NR, and Accuracy of Items Concerning "wear"

	Both OK	Both OK	Both OK	Only L2	Only L2	Only L2			
	wear coat	wear jeans	wear shoes	wear glasses	wear earrings	wear necktie			
C	19	18	15	18	10	14			
	95%	90%	75%	90%	50%	70%			
E	1	2	4	2	7	6			
	5%	10%	20%	10%	35%	30%			
NR	0	0	1	0	3	0			
	0%	0%	5%	0%	15%	0%			
	穿 外套	穿 牛仔裤	穿 鞋	戴 眼镜	戴 耳环	系 领带	穿眼镜*	穿耳环*	穿领带*
C	20	20	19	20	19	19	17	13	11
	100%	100%	95%	100%	95%	95%	85%	65%	55%
E	0	0	1	0	1	1	3	6	7
	0%	0%	5%	0%	5%	5%	15%	30%	35%
NR	0	0	0	0	0	0	0	1	2
	0%	0%	0%	0%	0%	0%	0%	5%	10%
	keep coat*	keep jeans*	keep shoes*	keep glasses*	keep earrings*	keep necktie*			
C	12	9	9	9	9	10			
	60%	45%	45%	45%	45%	50%			
E	7	10	10	11	8	7			
	35%	50%	50%	55%	40%	35%			
NR	1	1	1	0	3	3			
	5%	5%	5%	0%	15%	15%			

As the really high accuracy, the average RT of only the correct response data of the items

showed in Table 29 below was analyzed. The average RT of all English items is significantly longer than equivalent Chinese items (t (5) = 3.5251, $p<.05$), while for the three English items which have equivalent distractor Chinese items, there is no significant difference between English and Chinese items. The average RT of equivalent Chinese items and distractor Chinese items is not significantly different, too. The similar RT of "wear coat" and "wear glasses" may mean that the equivalence of "wear" is not only "穿", but both "穿" and "戴" in Chinese.

Table 29

The Average RT Data of Some Items Concerning "wear"

	wear + coat	wear + jeans	wear + shoes	wear + glasses	wear + earrings	wear + necktie
Average RT	1294	1241	1495	1220	1778	1841
	穿 + 外套	穿 + 牛仔裤	穿 + 鞋	戴 + 眼镜	戴 + 耳环	系 + 领带
Average RT	1195	1019	940	1062	1124	1073
				穿 + 眼镜	穿 + 耳环	穿 + 领带
				*	*	*
Average RT				1144	1396	1314

According to the results and the average data above, in *Figure* 21, the network of L1 and L2 taking "wear‧穿 (coat, jeans, shoes, glasses, earrings, and necktie)" as examples is depicted. The intralinguistic links are showed by black solid lines with RT (a shorter RT means a strong link reflected by a stronger solid line, while a longer RT means a weak link reflected by a thinner one). The interlinguistic links are showed by dotted lines. Some inappropriate associations which negative evidence is considered necessary are showed by red lines. The rectangles or ellipses with the same color means equivalence of two words in two languages. "keep" as the distractor of "wear" is showed in red font.

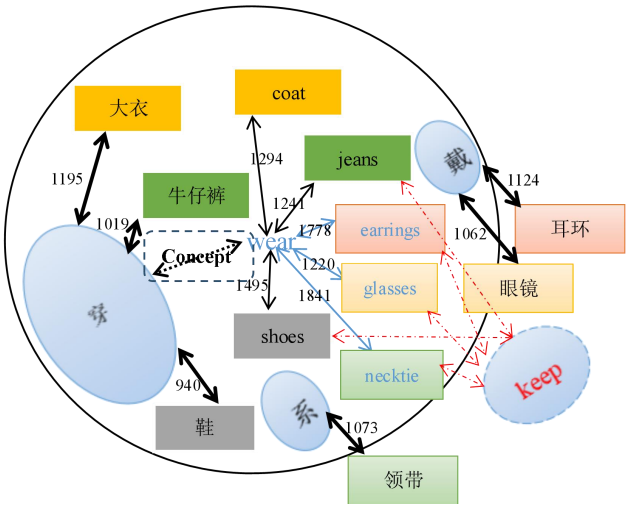


Figure 21.

The network of L1 and L2 taking "wear·穿 (coat, jeans, shoes, glasses, earrings, and necktie)" as examples.

5.4.1.4 break (L1<L2)

As for the results in the judgment test about "break", similar to the result of "wear", the difference among all English items and all Chinese items seemed small, too. All the items indicated relatively high accuracy. The accuracy of items about "break" in judgment test can be found in Table 30.

For items "bring + glasses*" and "bring + camera*", which were supposed to use as distractor items of "break", they can collocate to express totally different meanings from "break" actually, the two items are considered as problem items of experiment design and excluded from analysis.

Table 30

The Results of C, E, NR, and Accuracy of Items Concerning "break"

break+	glasses	window	record	tradition	camera	rule	promise	heart
C	15	17	16	16	12	17	16	18
	75%	85%	80%	80%	60%	85%	80%	90%
E	3	3	4	3	6	3	4	2
	15%	15%	20%	15%	30%	15%	20%	10%
NR	2	0	0	1	2	0	0	0
	10%	0%	0%	5%	10%	0%	0%	0%
	打碎+玻璃	打破+窗户	打破+纪录	打破+传统	弄坏+照相机	违反+规则	违背+承诺	
C	19	19	19	19	20	19	20	
	95%	95%	95%	95%	100%	95%	100%	
E	1	1	1	0	0	1	0	
	5%	5%	5%	0%	0%	5%	0%	
NR	0	0	0	1	0	0	0	
	0%	0%	0%	5%	0%	0%	0%	
	违背+玻璃*	违反+窗户*	打碎+纪录*	弄坏+传统*	违反+照相机*	打碎+规则*	弄坏+承诺*	
C	18	20	16	17	20	14	20	
	90%	100%	80%	85%	100%	70%	100%	
E	0	0	4	2	0	5	0	
	0%	0%	20%	10%	0%	25%	0%	
NR	2	0	0	1	0	1	0	
	10%	0%	0%	5%	0%	5%	0%	
bring+	~~glasses~~*	window*	record*	tradition*	~~camera~~*	rule*	promise*	heart*
C		17	12	15		13	13	15
		85%	60%	75%		65%	65%	75%
E		3	6	3		6	5	3
		15%	30%	15%		30%	25%	15%
NR		0	2	2		1	2	2
		0%	10%	10%		5%	10%	10%

As the really high accuracy, the average RT of only the correct response data of the items showed in Table 31 was analyzed. The average RT of English items is significantly longer than that of equivalent Chinese items (t (6) = 6.3574, $p<.05$) and the distractor Chinese items (t (6) = 3.4617, $p<.05$), while the average RT of two kinds of Chinese items (equivalence and distractor) items is not significantly different.

Table 31

The Average RT Data of Some Items Concerning "break"

	Average RT		Average RT			Average RT
break + glasses	1383	打碎 + 玻璃	1132	违背 + 玻璃	*	1290
break + window	1333	打破 + 窗户	1212	违反 + 窗户	*	1170
break + record	1595	打破 + 纪录	994	打碎 + 纪录	*	1280
break + tradition	1710	打破 + 传统	1204	弄坏 + 传统	*	1443
break + camera	1828	弄坏 + 照相机	1395	违反 + 照相机	*	1066
break + rule	1577	违反 + 规则	1203	打碎 + 规则	*	1428
break + promise	1700	违背 + 承诺	1191	弄坏 + 承诺	*	1402

According to the results and the average data above, the network of L1 and L2 taking "break (glasses, window, record, tradition, camera, rule, and promise)" as examples is depicted in *Figure* 22. The intralinguistic links are showed by black solid lines, and with a RT (a shorter RT = a stronger solid line). The interlinguistic links are showed by dotted lines. Some inappropriate associations which negative evidence is considered necessary are showed by red lines. The rectangles or ellipses with the same color means equivalence of two words in two languages. The distractor "bring" is showed in red font.

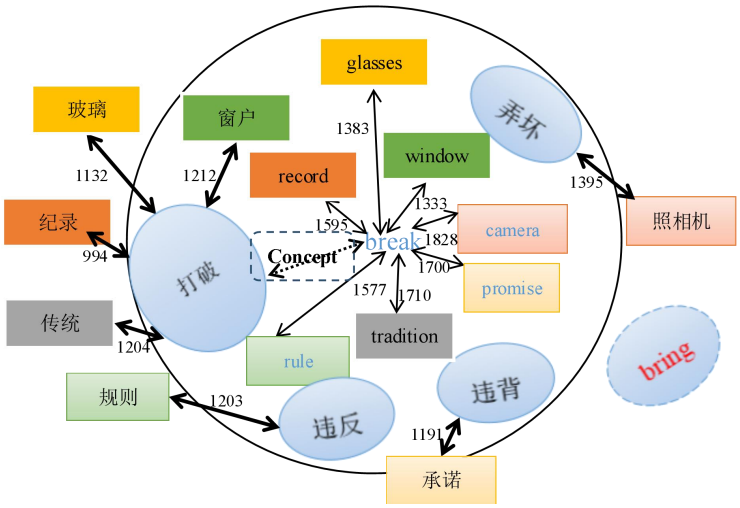


Figure 22.

The network of L1 and L2 taking “break as examples.

5.4.2 Cross-language priming lexical decision task

5.4.2.1 The Analysis of Reaction Times (RTs) and Error Rates (ERs)

Reaction times (RTs) and errors (Es) were recorded by the Super Lab automatically.

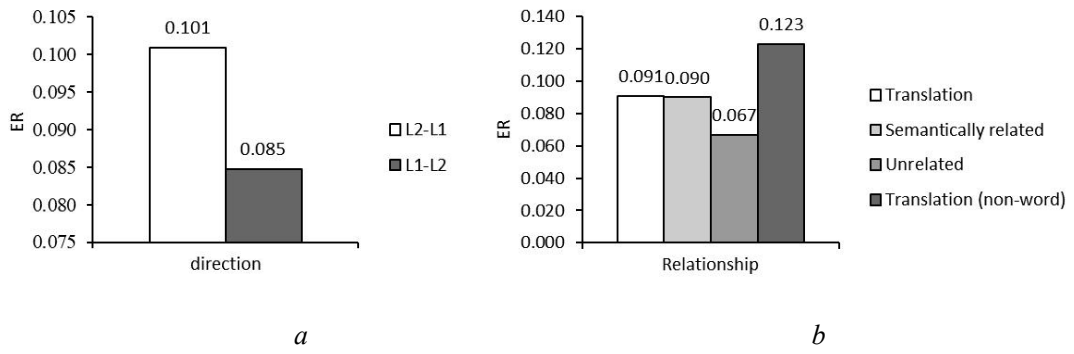
The number of correct (C) responses, errors, no response (NR) items, the accuracy / error rates (ERs) are presented in Table 32 below.

Table 32

The Results of C, E, NR, and the Accuracy / Error Rates (ER) of Lexical Decision Task in both Language Directions

	L1-L2					L2-L1				
	C	E	NR	AR	ER	C	E	NR	AR	ER
1	151	5	0	97%	3%	142	14	0	91%	9%
2	148	7	1	95%	5%	158	6	0	96%	4%
3	146	10	0	94%	6%	154	2	0	99%	1%
4	144	10	2	92%	8%	148	8	0	95%	5%
5	148	6	2	95%	5%	116	30	10	74%	26%
6	141	13	2	90%	10%	125	29	2	80%	20%
7	134	21	1	86%	14%	148	8	0	95%	5%
8	130	26	0	83%	17%	120	36	0	77%	23%
9	148	6	2	95%	5%	155	1	0	99%	1%
10	146	10	0	94%	6%	150	6	0	96%	4%
11	135	21	0	87%	13%	124	32	0	79%	21%
12	135	21	0	87%	13%	145	11	0	93%	7%
13	138	17	1	88%	12%	153	3	0	98%	2%
14	153	3	0	98%	2%	152	4	0	97%	3%
15	145	10	1	93%	7%	151	5	0	97%	3%
16	136	18	2	87%	13%	146	9	1	94%	6%
17	116	38	2	74%	26%	141	12	3	90%	10%
18	128	28	0	82%	18%	129	26	1	83%	17%
19	128	24	4	82%	18%	150	6	0	96%	4%
20	120	35	1	77%	23%	127	28	1	81%	19%

Within-subjects ANOVA conducted on the error rates revealed that the interaction of direction and pairs' relationship is not but almost significant ($F(3, 57) = 2.818$, $p = .052$, $\eta_p^2 = .129$), and the main effects of direction and pairs' relationship are both not significant.



a *b*

Figure 23.

The effects of language direction (a) and pairs' relationship (b) in lexical decision task.

Then, RTs for correct responses on the lexical decision task were analyzed. Task performance was highly accurate in all conditions (average 91%), so only the RTs to word targets were submitted to within-subjects t test and ANOVA.

Table 33

The RT to Targets with Different Relationships to Stimulus in Lexical Decision Task in both Directions

	L2-L1				L1-L2			
	Translation	Related	Unrelated	Translation (non-words)	Translation	Related	Unrelated	Translation (non-words)
1	907	1029	1110	918	896	990	964	922
2	524	557	577	599	607	614	633	678
3	602	840	688	758	786	841	874	879
4	737	770	774	1089	897	927	938	1007
5	1081	1249	1111	1277	957	1056	1039	972
6	797	1107	1237	854	692	732	703	908
7	735	752	666	914	714	681	897	1084
8	425	445	460	475	488	540	537	552
9	782	793	686	914	904	932	1089	1093
10	737	746	805	810	786	790	774	895
11	569	563	575	600	517	543	570	661
12	570	606	567	688	754	828	1049	953
13	574	676	612	687	781	819	881	945
14	502	540	541	543	552	540	560	623
15	725	815	669	826	739	836	825	908
16	860	873	836	869	954	860	843	952
17	1035	1168	1096	984	1159	1070	1155	1418
18	819	784	699	722	899	714	789	735
19	626	580	581	711	643	813	850	1194
20	523	459	522	489	603	510	542	541

First, within-subjects t tests were used to confirm the translation priming effect. Results showed that there is no significant difference between translation pairs and unrelated pairs in L2-L1 direction, but a faster RT of translation pairs than unrelated pairs in L1-L2 direction ($t(19) = -2.581$, $p < .05$, *Figure* 24), though the $r = -.171$ and $d = -0.331$ are very low.

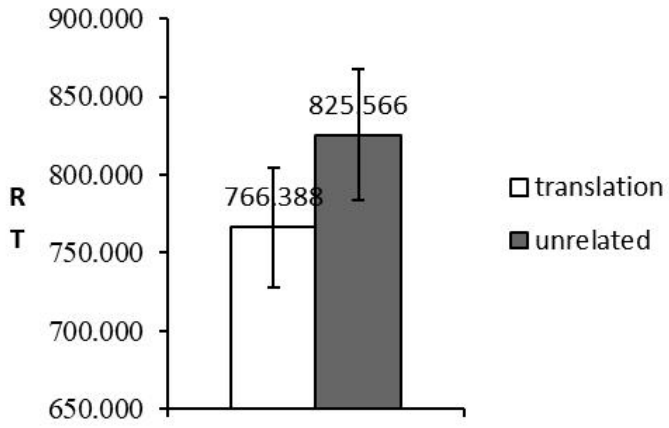


Figure 24.

The average RT of translation pairs and unrelated pairs in L1-L2 direction.

Then, within-subjects ANOVA was conducted as within-subject factors were language direction (L1-L2, L2-L1), and relationship (translation, related, unrelated, and translation pairs with non-word targets). Results showed that the interaction of direction and pairs' relationship is not significant ($F(3, 57)=2.641$, $p=.071$, $\eta_p^2=.122$), however, the main effects of direction ($F(1, 19)=5.549$, $p=.029<.05$, $\eta_p^2=.226$) and pairs' relationship ($F(3, 57)=11.231$, $p<.001$, $\eta_p^2=.372$) are both significant, with significantly faster RTs to L1 targets than to L2 targets, and significantly faster RTs to correct targets preceded by a translation, and related prime than to targets following an unrelated prime, or incorrect targets preceded by a translation (see Table 33 and *Figure* 25).

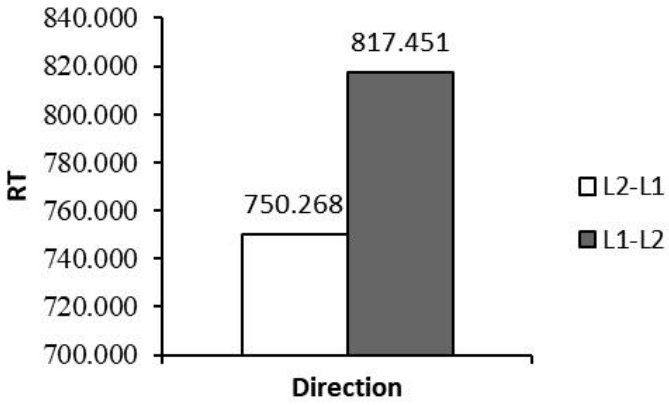


Figure 25.

The average RT in L2-L1 direction and L1-L2 direction.

The analysis of simple main effect of direction showed that for the pairs' relationship of unrelated, and translation pairs with non-word targets, the RTs of two directions are significantly different (F (1, 76) =5.618, p=.020<.05, η_p^2=.228; F (1, 76) =9.357, p=.003<.05, η_p^2=.330) with L2-L1 is faster than L1-L2, while for the translation and related pairs, the RTs' difference of two directions are not significant.

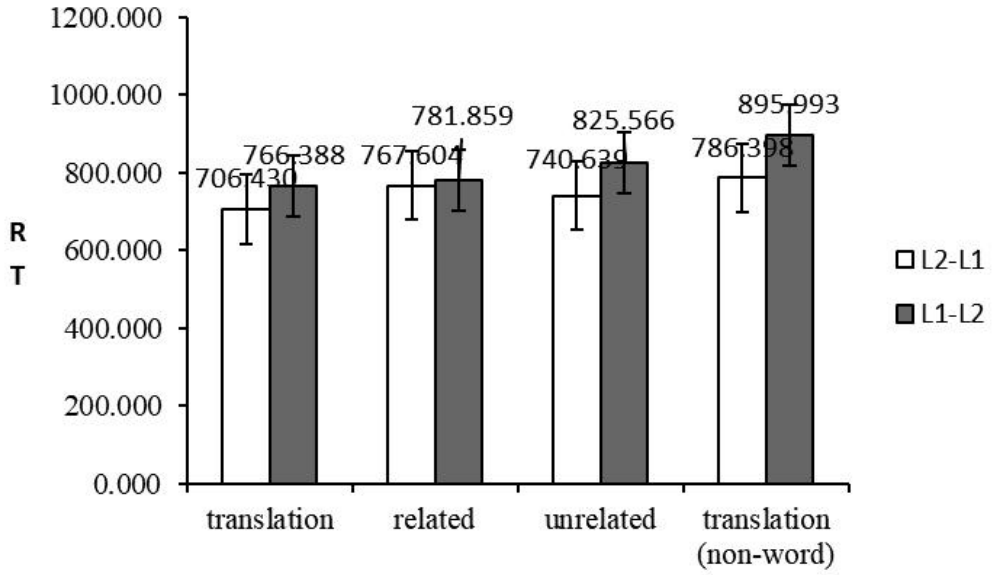


Figure 26.

The difference of average RT between two directions in each kind of pairs' relationship.

The analysis of simple main effect of pairs' relationship showed that for the both directions, RTs of are significantly different (F (3, 14) =3.729, p=.021<.05, η_p^2=.228; F (3, 14) =10.411, p=.000<.001, η_p^2=.354). Especially for the direction L1-L2, the RT of relationship translation is significantly faster than translation pairs with non-word targets (adjusted p=.005<.05); and related is faster than unrelated (adjusted p=.019<.05), and translation pairs with non-word targets (adjusted p=.006<.05) (See *Figure 27* below).

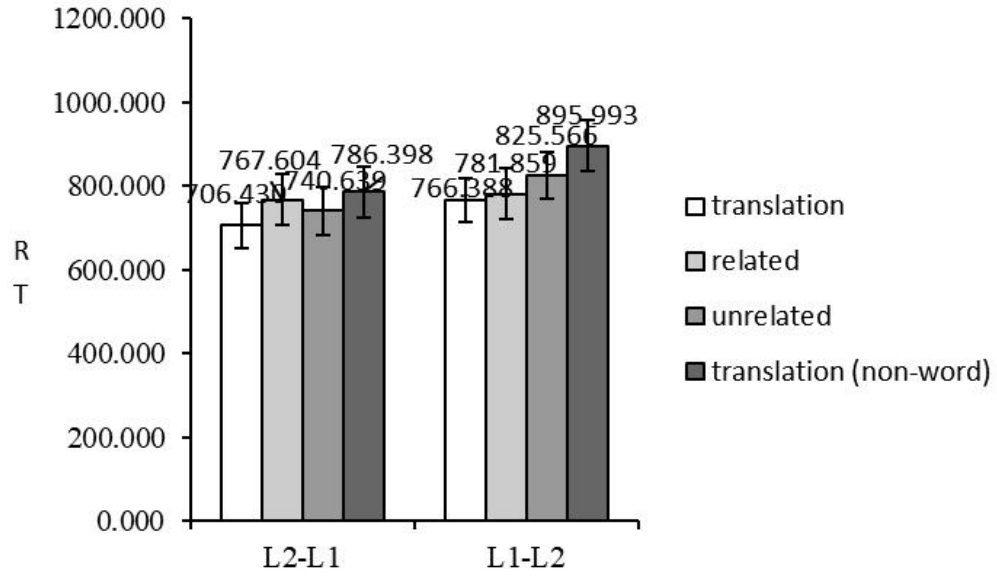


Figure 27.

The difference of average RT among pairs' relationship in each direction.

5.4.2.2 The Analysis of Translation Pairs of Different Semantic Category

RTs for correct responses on the translation pairs with both direction (L1-L2 and L2-L1) were submitted to a within-subjects ANOVA. Within-subjects factors were language direction (L1-L2, L2-L1), and category relationship (L1>L2, L2<L1). Results showed that the interaction of direction and category relationship is significant (F (1, 19) =18.021, p<.000, η_p^2=.487), however, the main effects of direction (F (1, 19) =0.692, p=.416, η_p^2=.035) and category relationship (F (1, 19) =1.109, p=.305, η_p^2=.055) are both not significant.

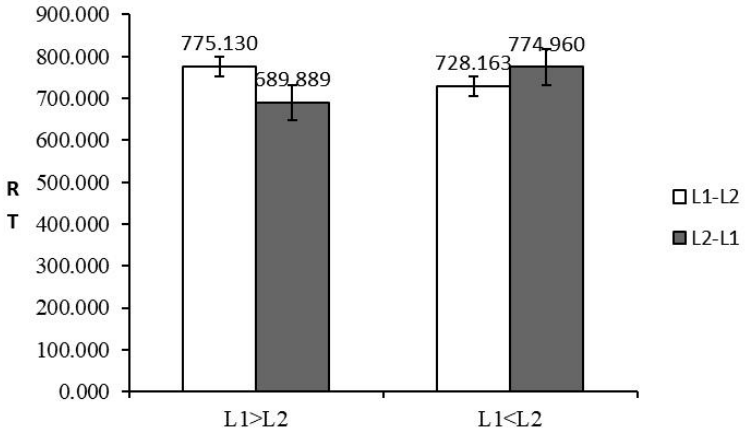


Figure 28.

The effects of language directions for different category relationship.

The analysis of simple main effect of direction showed that for the semantic category relationship L1>L2, the RTs of two directions are significantly different (F (1, 38) =9.370, p=.004<.05, η_p^2=.330) with L2-L1 is faster than L1-L2, while for the semantic category relationship L1<L2, the RTs' difference of two directions are not significant.

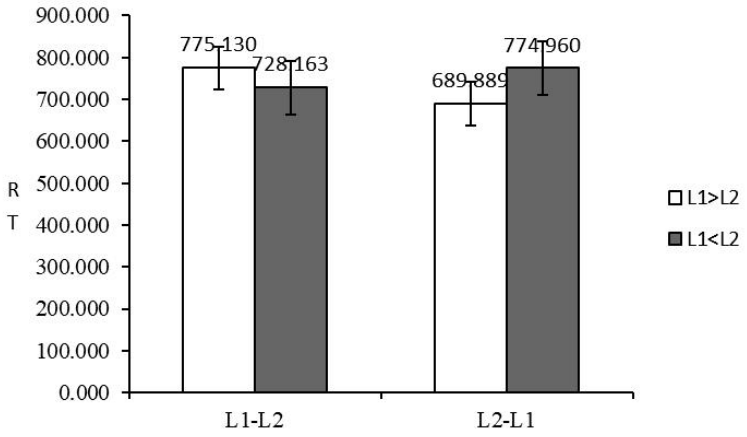


Figure 29.

The effects of category relationships in different language directions.

The analysis of simple main effect of semantic category relationship showed that for the direction L2-L1, the RTs of two kinds of semantic category relationship are significantly different (F (1, 38) =12.716, p=.001<.05, η_p^2=.401) with L1<L2 is faster than L1>L2, while for the direction L1-L2, the RTs' difference of two kinds of semantic category relationship are not significant.

5.4.3 Comparing the HIGH and LOW VST Score Groups

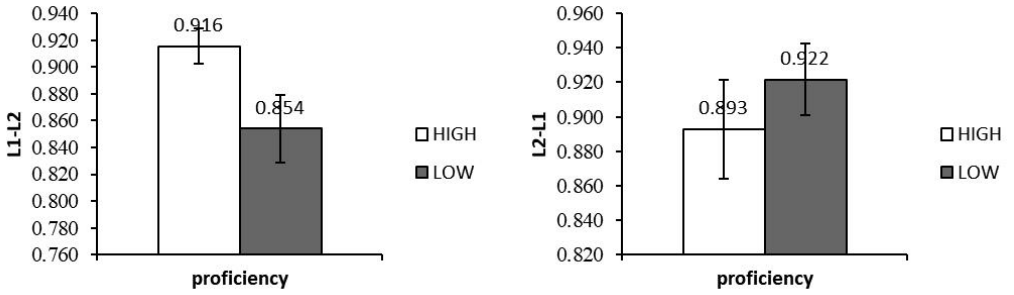
As the proficiency of participants was evaluated by the VST scores. The average score was used as a criterion of either it is HIGH or LOW.

The average score of their VST was 59.65. And as mentioned in Experiment 2, according to Nation's (2006) estimates, and as supported in several subsequent studies, 6,000 to 9,000 word families can account for 98% of text coverage, which, in turn, enables independent reading.

Participants whose VST is equal or greater than 60 is evaluated as HIGH and lower than 60 as LOW. In other words, participants with a HIGH vocabulary size are considered as an independent reader in L2 English, while participants with a LOW vocabulary size are considered do not have enough vocabulary to read L2 materials independently. According to the level of participants' vocabulary size evaluated by the VST, there were 11 participants with HIGH proficiency and 9 with LOW.

Welch's t tests were performed to determine if there was a statistically significant difference in the accuracy and the RTs of translation priming between HIGH proficiency and that of LOW proficiency participants. A *p* value less than 0.05 was considered statistically significant.

Results showed that in L1-L2 direction, the accuracy of HIGH proficiency participants is significantly higher than that of LOW proficiency participants (t (12.414) =2.175, $p<.05$). Although the r=.475 and d=0.986 are a little small. In L2-L1 direction there was no significant difference between two proficiency groups (t (17.297) =-0.815, $p> .05$).

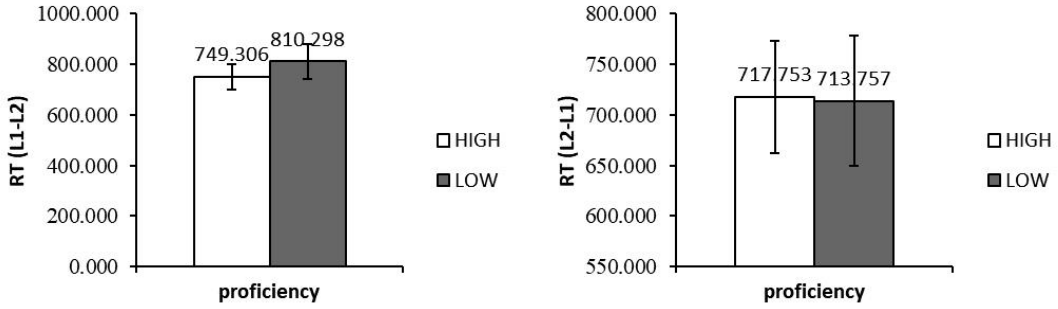


a *b*

Figure 30.

The comparison of accuracy between HIGH and LOW groups in L1-L2 direction (a) and L2-L1 direction (b).

As for RTs, results showed that both in L1-L2 direction, and in L2-L1 direction, the RTs of HIGH proficiency participants and that of LOW proficiency participants are not significantly different (L1-L2 t (13.607) =-0.735, $p>.05$; L2-L1 t (15.393) =0.047, $p>.05$).



a *b*

Figure 31.

The comparison of RT between HIGH and LOW groups in L1-L2 direction (a) and L2-L1 direction (b).

In conclusion, in both L1-L2 and L2-L1 directions, the HIGH and LOW proficiency participants response to targets in a similar length of time in both directions, and have a similar accuracy in L2-L1 direction, while HIGH proficiency participants showed a significantly higher accuracy in L1-L2 direction.

5.5 General Discussion

5.5.1 About the Effect of Semantic Category Relationship

The results of accuracy in collocation decision task were different with the different relationship between the semantic categories of L1 and L2. The difference between association words concerning “drink” (“water”, “soup”, “medicine”), “open” (“door”, “TV”, “letter”) which in L1 Chinese have a broader semantic category and “wear” (“coat”, “jeans”, “shoes”), “break” which in L2 English have a broader semantic category verified Bi (2020)’s and Experiment 1’s proposition that words like “drink” which in L1 Chinese have a broader semantic category are considered to be more difficult to acquire than words like “wear” and “break” which in L2 English have a broader semantic category.

Specifically, for Chinese participants, the accuracy of “drink soup*”, “drink medicine*”, and “eat medicine*” etc. which expressed in a Chinese way as “喝汤” (drink soup*), “喝药” (drink medicine*) and “吃药” (eat medicine*) and needed to be rejected was really low comparing to others. It is still not able to reject the wrong expression in English completely for not only the low

proficiency learners but also the high proficiency learners. These evidence means that these collocations are difficult to be acquired as suggested in Bi (2020). Additionally, in this study we newly know that even the learners cannot be able to reject the wrong expressions completely, they have accepted some of the correct expressions in the meantime. In other words, both wrong expressions affected by L1 and correct L2 expressions coexist in some learners' mental lexicon.

On the other hand, the differences among all items of "wear" and "break" which L2 English have a broader semantic category were not that apparent. The accuracy of most items was high. This means that these collocations are acquired easily. Additional discovery besides the results of previous study, is that even learners are able to acquire the correct expressions easily, the distractor words "keep" or "bring" may affect somehow at the same time. Especially "keep" which have a general and comprehensive meaning in English is considered to affect the lexicon of "wear 穿"to some extent. However, as there were no pictures or videos showed in this experiment, it is difficult to attribute this effect to whether participants misjudged items concerning "keep" because they considering "keep" as a word with similar meaning as "wear" or just misjudged the items in terms of a totally different meaning from "wear" of "keep". The details need to be studied in further research.

In one word, these results of accuracy confirmed the mediation of L1 vocabulary knowledge in the construction of L2 lexical network, and indicated that the influence of L1 word knowledge to word association was different with the different relationship between the semantic categories of L1 and L2.

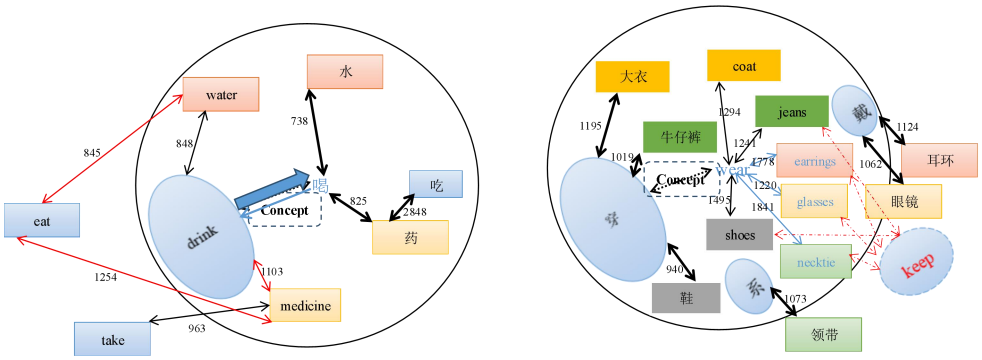
Moreover, the RT data in this study provided some new evidence about lexical network. For "喝"("drink") and "打开"("open") which L1 Chinese has a broader semantic category, the low accuracy data of some items makes the analysis of RT complicated, because the RT of CORRECR response and ERROR response should not be treated in the same way or be compared directly. The average data are not representative in some cases, therefore, the network of representation of participants for every typical situation were depicted. It seems that although not only the low proficiency learners but also the high proficiency learners were still not able to reject the wrong expression completely; some learners have the correct L2 links in their mental lexicon while others do not. What calls for special attention is that whether the effect of proficiency makes this difference need to be studied statically with more data. On the other hand, for "wear" and "break"

which L2 English has a broader semantic category, the average RT of correct responses reflected that the links between two words in Chinese items is significantly stronger than words in equivalent English items.

In the lexical decision task, for the semantic category relationship L1>L2, the RTs in L2-L1 direction is faster than in L1-L2 direction, while for the semantic category relationship L1<L2, the RTs' difference of two directions are not significant.

For the semantic category relationship L1>L2, the asymmetry of RTs in two directions may indicates that the L1 word 喝 with larger semantic category equals to more than one L2 words (drink and take, for example), so the activation of the special one (drink) implemented after retrieving among many words and then reducing to the just one the same as the target. One the other hand, in L2-L1 direction, participants need to just retrieve the only equivalence of L2 "drink" in L1 as 喝. Therefore, the process takes less time than in L1-L2 direction. However, for the semantic category relationship L1<L2, the inverse results did not appear significantly. The number of items may affect to some extent, as the number of items with semantic category relationship L1<L2 is much less than L1>L2 in this experiment. Further, this may because when the L2 word "wear" with larger semantic category equals to more than one L1 words (穿, 戴, and 系, for example), the retrieval and reduce process occur similarly as for the semantic category relationship L1>L2. One the other hand, in L1-L2 direction, participants only need to retrieve the equivalence of L1 in L2, similarly. For 戴, with typical equivalence relationship with "wear" it may be easy. However, for 戴, for example, the not that typical equivalence relationship may take time to recall "wear" in L2. There may be some difference among items.

To put all together, *Figure* 19 and *Figure* 21 can be depicted further detailed in links direction as showing in Figure 32 with a thick arrow from L2 drink to L1 喝 and a thin arrow from L1 喝 to L2 drink; while with a similar thickness between L1 穿 and L2 wear.



L1>L2 L1<L2

Figure 32. (Modification of *Figure 19 & 21*)

The network of L1 and L2 taking "drink‧喝 and wear‧穿" as examples with information of links direction between L1 and L2.

5.5.2 About the Translation Priming

In the lexical decision task, the results of accuracy are not significantly different in both L1-L2 and L2-L1 directions and among different pairs. However, RTs in L2-L1 direction is significantly faster than in L1-L2 direction, and RTs to correct targets preceded by a translation, and related prime are significantly faster than to targets following an unrelated prime, or incorrect targets preceded by a translation. A weak translation priming effect is appeared in L1-L2 direction.

There was an asymmetric cross-language priming: A significant but weak translation-priming effect and cross-language semantic priming were observed in only L1-L2 direction. This result agrees with many previous studies have observed translation priming only in the L1-L2 direction and cross-language semantic priming (Referring **2.7** for lists of previous studies).

These findings fit well with two kinds of mechanisms which has introduced in the previous studies (Referring **2.1.2** and **2.1.3**). Both of the two models mentioned the reduced efficiency of L2 primes.

The RHM (Kroll & Stewart, 1994; Kroll, van Hell, Tokowicz, & Green, 2010. Referring **2.1.2**) argues that a low conceptual access efficiency from L2 primes lead to a reduced priming in the L2-L1 direction. This model proposes that L2 learners' the lexical forms of two languages are represented separately, but just represented commonly in conceptual representations. At the conceptual level, the connection between representations and concepts in L1 is stronger than in L2. Initially, L2 learners access the meaning of L2 words through L1 translation equivalents, until

their L2 proficiency approve. According to the revised hierarchical model, L1 primes activate the conceptual representations shared with L2 targets in L1-L2 direction, facilitating recognition of L2 targets. For learners with not high enough proficiency, the connection between L2 representations and concepts are weak, therefore it is difficult for concepts to be activated in L2-L1 direction.

By prolonging prime presentation duration in a masked priming lexical decision task, priming effects may become strong, for the providence of much more time, the weak connection between L2 representations and their concepts may be activated. This means that the priming asymmetry is a result of differences in the difficulty and speed of access.

While what highlighted by the BIA+ model (Referring **2.1.3**) is focusing on the resting-level activation of both the primes and the targets. In L1-L2 direction, a strong L1 prime with its higher resting-level activation led to a more effective prime effect to the L2 targets whose resting-level activation was lower and might benefit from the prime to a larger extent. On the other hand, in the opposite direction L2-L1, the combined effects of a lower resting-level activation of the L2 prime and a higher resting-level activation of the L1 target resulted in the absence of any facilitation, as has been reported in several previous studies (e.g., Basnight-Brown & Altarriba, 2007; Smith, Y., Walters, J., & Prior, A., 2019).

It has been suggested in the past that translation priming occurs because such pairs are connected at the lexical level and that semantics and conceptual representations are not activated when translation equivalents are processed (de Groot & Nas, 1991). However, in this study both translation priming and semantic related priming in L2-L1 direction is significant. This indicates that participants rely solely on a lexical level of processing, or activate semantics and conceptual representations in both translation priming and semantic related priming in this process in L2-L1 direction.

In comparison to the results of the translation priming effect and the cross-language semantic priming effect, it failed to reach significance, although some findings like that there was no significant cross-language semantic priming (e.g., Basnight-Brown & Altarriba, 2007; de Groot & Nas, 1991) or it was much weaker (Smith, Y., Walters, J., & Prior, A., 2019) have been reported in several previous studies. The range of meaning overlap of semantic related pairs is considered to be smaller and more variable than translation pairs in those studies. However, it is not verified in this study.

The two results above implicated a possibility of processing in solely lexical level, rather than semantics and conceptual representations level in both translation priming and semantic related priming in L2-L1 direction. In other words, semantic related priming is just nominal here and the process is on lexical level only. This indicates that the L2 network are formed on the basis of L1 and may could provide an explanation of the usually lower stability of L2 network compared with L1.

5.5.3 About the Effect of L2 Proficiency

In both L1-L2 and L2-L1 directions, the HIGH and LOW proficiency participants have a similar accuracy in L2-L1 direction, while HIGH proficiency participants showed a significantly higher accuracy in L1-L2 direction. All participants response to translation targets with similar RTs. In other words, this means that in L1-L2 direction, in an almost equally length of time, the L2 recognition may influenced by the learners' L2 proficiency to some extent. The HIGH proficiency participants may have a solid foundation of L2 lexical network than the LOW proficient participants. This can also be explained by RHM (Kroll & Stewart, 1994; Kroll, van Hell, Tokowicz, & Green, 2010) and by the results of collocation decision tasks to some extent. LOW proficiency L2 learners access the meaning of L2 words through L1 translation equivalents, or they were on the way to access the meaning of L2 words without support of L1 translation equivalents, so there may be a situation in which more incorrect links in L2 which misguided by L1 are stay in their lexical network. Therefore, LOW proficiency learners' network is less stable than HIGH proficiency participants and may appear more mistakes in a limited time.

5.6 Conclusion

The aim of the present study was to gain insight into the way in which Chinese English learners represent their two languages in mental lexicon: especially about how the words linked with each other from the aspect of direction and strength. A collocation decision task was implemented to investigate the collocational links within language and to compare the equivalent collocations between language, and to infer whether or not and how the semantic category of L2 learners' L1 Chinese influence their L2 English vocabulary learning. An experimental priming paradigm under masked condition was used to examine how semantically related and translation

word pairs would be recognized for L2 learners. Based on Experiment 1, collocations concerning “drink”, “open”, which in L1 Chinese have a broader semantic category; and “wear”, “break”, which in L2 English have a broader semantic category were used in collocation decision task, while more words which have either the kind of relation between L1 and L2 semantic category (L1>L2 or L1<L2) were selected as items for lexical decisions task. Twenty adult Chinese learners of English participated in this experiment by a remote way online. The accuracy and reaction time (RT) data collected in the tasks were analyzed.

5.6.1 Summary of the Findings

The results of accuracy in collocation decision task confirmed the mediation of L1 vocabulary knowledge in construction of L2 lexical network, and indicated that the influence of L1 vocabulary knowledge to L2 lexical network is distinguishing by different relationships between the semantic categories of L1 and L2. According to the results of RTs which provided some information about the strength of links between two words in a collocation, learners’ lexical networks of some representative words were depicted more detailed in Experiment 3 on the basis of Experiment 1.

In the lexical decision task, for the semantic category relationship L1>L2, the RTs in L2-L1 direction is faster than in L1-L2 direction; while for the semantic category relationship L1<L2, the RTs’ difference of two directions are not significant. The results of accuracy are not significantly different in both L1-L2 and L2-L1 directions and among all kinds of pairs. However, RTs in L2-L1 direction is significantly faster than in L1-L2 direction. A weak translation priming effect is appeared in L1-L2 direction.

As for the L2 proficiency effect, in both L1-L2 and L2-L1 directions, the HIGH and LOW proficiency participants have a similar accuracy in L2-L1 direction, while HIGH proficiency participants showed a significantly higher accuracy in L1-L2 direction. All participants response to translation targets with similar RTs. These results are discussed with regard to typical models of bilingual memory representation (RHM & BIA+).

5.6.2 Limitations and Recommendations for Further Research

There are two limitations that need to pay attention to in further research.

First, in the collocation decision task, the collocation is simply presented as “A+B” this time. However, as the prevalent polysemy in language, in addition to the common meaning, some participants know the rare meaning of some words and the decision become different, therefore, to limit the meaning to an extremely exact point which is going to be discussed, it is better to provide some pictures or videos additionally, besides literal presentation in further studies.

Second, in lexical decision task, semantic related pairs were confirmed by the writer and other three people to make sure that they were actually related semantically. However, how the targets related to the stimulus words (synonym, antonym, hypernym, hyponym, or etc.) may affect the reaction times and should be consider more carefully in future research.

To briefly sum up all three experiments in this study, as one of the most important findings, both associated words in word association tasks in Experiment 1 and the accuracy and RT in collocation decision tasks in Experiment 3 provided the evidence that L1 word knowledge affect L2 vocabulary recognition and acquisition in some different way in light of the range relationship of semantic category between L1 and L2. Even the mental lexicon seems highly complex and further far from known-well, this study attempted to portray lexical network of L2 learners taking representative words as examples. The direction and strength information of links within language and interlanguage in L2 learners’ mental lexicon become clear in the network figure. Then, the influence of L2 vocabulary knowledge (learners L2 proficiency, generally speaking) to L2 word association in Experiment 2 and to lexical decision in Experiment 3 were also confirmed in the present study. In a word, this study took advantage of Word Association Task and online Recognition Task and verified the effects of semantic category relationship between L1 and L2 and L2 vocabulary knowledge in L1 and L2 lexical network of Chinese EFL learners.

Results of this study will provide evidence about L1 and L2 lexical network for L2 mental lexicon research area and give support to semantic transfer in L2 learning in the SLA research area. Then, it is expected that this study could promote further research on words’ semantic category or on vocabulary semantic analysis from the perspective of linguistics, which may offer new orientations to research of vocabulary recognition in mental lexicon and vocabulary acquisition research in return. Furthermore, this study may give some pedagogical enlightenment to education field as words with different types of semantic category relationship between L1 and L2 may need

to be taught in different ways.

REFERENCES

Aitchison, J. (2012). *Words in the mind: An introduction to the mental lexicon*, 4th Edition. Chichester, UK: John Wiley & Sons.

Akmajian, A., & Demers, R. A., & Farmer, A. K., & Harnish, R. M. (2010). *Linguistics An Introduction to Language and Communication*, Sixth Edition. MIT.

Altarriba, J. (1992). *The representation of translation equivalents in bilingual memory.* In R. J. Harris (Ed.), Cognitive processing in bilinguals. Amsterdam: Elsevier. pp. 157-174

Altarriba, J., & Basnight-Brown, D. M. (2007). Methodological considerations in performing semantic- and translation-priming experiments across languages. *Behavior Research Methods, 39,* 1-18.

Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System 21(1),* 101-114.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception & Performance, 10,* 340-357.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory & Language, 24,* 89-106.

Basnight-Brow, D.M, & Altarriba, J. (2007). Differences in semantic and translation priming across languages: The role of language direction and language dominance. *Memory & Cognition, 35(5),* 953-965.

Batty, A. (2012). Identifying dimensions of vocabulary knowledge in the word associates test. *Vocabulary Learning and Instruction, 1(1),* 70-77.

Bennett, P., & Stoeckel, T. (2012). Variations in format and willingness to skip items in a multiple-choice vocabulary test. *Vocabulary Education and Research Bulletin, 1,* 2-3.

畢奇 (2020). L1 意味カテゴリーが L2 語彙習得に与える影響―日本語母語話者および中国語母語話者である英語学習者を対象にして―, *電子情報通信学会研究報告, TL2020-2,* 6-11.

Chen, H.C., & Ng, M.L. (1989). Semantic facilitation and translation priming effects in Chinese–English bilinguals. *Memory & Cognition, 17,* 454-462.

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

Chumbley, J. I., & Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Memory & Cognition, 12,* 590-606.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82,* 407-428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning & Verbal Behavior, 8,* 240-247.

Cronbach, L. J. (1942). An analysis of techniques for diagnostic vocabulary testing. *The Journal of Educational Research, 36,* 206-217. https://doi.org/10.1080/00220671.1942.10881160

Cronin, V. S. (2002). The syntagmatic-paradigmatic shift and reading development. *Journal of Child Language, 29,* 189-204.

Cremer, M., Dingshoff, D., de Beer, M., & Schoonen, R. (2011). Do word associations assess word knowledge? A comparison of L1 and L2, child and adult word associations. *International Journal of Bilingualism, 15,* 187-204.

Dang, T. N. Y., & Webb, S. (2014). The lexical profile of academic spoken English. *English for Specific Purposes, 33,* 66-76.

De Deyne, S., Navarro, D. J., & Storms, G. (2012). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior Research Methods, 45(2),* 480-498.

De Groot, A. M. B. (1984). Primed lexical decision: Combined effects of the proportion of related prime–target pairs and the stimulus onset asynchrony of prime and target. *Quarterly Journal of Experimental Psychology, 36A,* 253-280.

De Groot, A. M. B. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(5),* 824-845.

De Groot, A. M. B. (1992). Determinants of Word Translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1001-1018.

De Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in

compound bilinguals. *Journal of Memory & Language, 30,* 90-123.

De Groot, A. M. B., & Poot, R. (1997). Word Translation at Three Levels of Proficiency in a Second Language: the Ubiquitous Involvement of Conceptual Memory. *Language Learning, 47,* 215-264.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,* 175-197.

Dronjic, V., & Helms-Park, R. (2014). Fixed-choice word-association tasks as second-language lexical tests: What native speaker performance reveals about their potential weaknesses. *Applied Psycholinguistics, 35(1),* 193-221.

Duyck, W., & Brysbaert, M. (2004). Forward and Backward Number Translation Requires Conceptual Mediation in Both Balanced and Unbalanced Bilinguals. *Journal of Experimental Psychology: Human Perception and Performance, 30(5),* 889-906. https://doi.org/10.1037/0096-1523.30.5.889

Duyck, W., & Brysbaert, M. (2008). Semantic access in number word translation: The role of crosslingual lexical similarity. *Experimental Psychology, 55(2),* 102-12. https://doi.org/10.1027/1618-3169.55.2.102

Entwisle, D. R., Forsyth, D. F., & Muuss, R. (1964). The syntactic-paradigmatic shift in children's word associations. *Journal of Verbal Learning and Verbal Behavior, 3(1),* 19-29.

Ervin, S. M. (1961). Changes with age in the verbal determinants of word-association. *The American Journal of Psychology, 74(3),* 361-372.

Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory & Language, 51,* 1-22.

Fitzpatrick, T. (2006). Habits and rabbits: Word associations and the L2 lexicon. *EUROSLA Yearbook, 6,* 121-145.

Fitzpatrick, T. (2007). Word association patterns: Unpacking the assumptions. *International Journal of Applied Linguistics, 17,* 319-331.

Fitzpatrick, T. (2009). Word association profiles in a first and second language: Puzzles and problems. in T. Fitzpatrick & A. Barfield (Eds.), *Lexical processing in second language learners* (pp. 38-52). Clevedon, UK: Multilingual Matters.

Fitzpatrick, T. (2012). Tracking the changes: vocabulary acquisition in the study abroad context. *The Language Learning Journal, 40(1),* 81-98.

Fitzpatric, T., & Izura, C. (2011). Word association in L1 and L2: An exploratory study of response types, response times, and interlingual mediation. *Studies in Second Language Acquisition, 33,* 373-398.

Fitzpatrick, T, & Thwaites, P. (2020). Word association research and the L2 lexicon. *Language Teaching, 53,* 237-274. doi:10.1017/S0261444820000105

Frenck-Mestre, C. & Prince, P. (1997). Second language autonomy. *Journal of Memory and Language, 37(4),* 481-501.

Frenck, C., & Pynte, J. (1987). Semantic representation and surface forms: A look at across-language priming in bilinguals. *Journal of Psycholinguistic Research, 16,* 383-396.

French, R. M. & Jacquet, M. (2004). Understanding bilingual memory: models and data. *TRENDS in Cognitive Sciences, 8(2),87-93.*

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: masked priming with cognates and non-cognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1122-1139.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58,* 787-814.

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General, 140,* 186-209.

Grainger, J., & Frenck-Mestre, C. (1998). Masked priming by translation equivalents in proficient bilinguals. *Language & Cognitive Processes, 13*, 601-623.

Grainger, J., & Beauvillain, C. (1988). Associative priming in bilinguals: Some limits of interlingual facilitation effects. *Canadian Journal of Psychology, 42,* 261-273.

Håkansson, G., & Norrby, C. (2010). Environmental influence on language acquisition: Comparing second and foreign language acquisition of Swedish. *Language Learning, 60(3),* 628-650.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language*

Acquisition, 21, 303-317.

Hiebert, E.H. & Kamil, M.L. (2005). *Teaching and learning vocabulary: bringing research to practice.* London: Lawrence Erlbaum Associate.

Higginbotham, G. (2010). Individual learner profiles from word association tests: The effect of word frequency. *System, 38(3),* 379-390.

Hutchison, K. A., Neely, J. H., & Johnson, J. D. (2001). With great expectations, can two "wrongs" prime a "right"? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27,* 1451-1463.

Imai, M. (1993). What is missing in L2 word meaning representation? Problems in second language vocabulary learning. *Japanese Journal of Educational Psychology, 41,* 243-253.

Izura, C., & Playfoot, D. (2012). A normative study of acronyms and acronym naming. *Behavior Research Methods, 44(3),* 862-889.

Jenkins, J. J. (1954). Transitional Organization: Association Techniques. in C. Osgood & T. Sebeok (Eds.), *Psycholinguistics: a survey of theory and research problems*. Baltimore: Waverly Press.

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.

Jiang, N. (2002). Form-meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition, 24,* 617-637.

Jiang, N. (2004). Semantic transfer and its implications for vocabulary teaching in a second language. *The Modern Language Journal, 88,* 416-432.

Jiang, N., & Forster, K. I. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory & Language, 44,* 32-51.

Jiang, N. & Zhang, J. Q. (2019). Form prominence in the L2 lexicon: Further evidence from word association. *Second Language Research, 00(0),* 1-22.

Keatley, C. W., & de Gelder, B. (1992). The bilingual primed lexical decision task: Cross-language priming disappears with speeded responses. *European Journal of Cognitive Psychology, 4,* 273-292.

Khazaeenezhad, B., & Alibabaee, A. (2013). Investigating the role of L2 language proficiency in word association behavior of L2 learners: a case of Iranian EFL learners. *Theory and Practice in Language Studies, 3(1),* 108-115.

Kikuchi, M., Yamamoto, M., Yoshimura, M., Yabuuchi, S., & Tanimura, M. (2001). Assessing the hierarchical structure of L2 mental lexicon: An analysis of categorical and non-categorical word association. *JACET Bulletin, 34,* 27-35.

Kiran, S., & Lebel, K. R. (2007). Crosslinguistic semantic and translation priming in normal bilingual individuals and bilingual aphasia. *Clinical Linguistics & Phonetics, 21(4), 277-303.* https://doi.org/10.1080/02699200701243873

Kirsner, K., Smith, M. C., Lockhart, R. S., King, M. L., & Jain, M. (1984). The bilingual lexicon: Language-specific units in an integrated network. *Journal of Verbal Learning & Verbal Behavior, 23,* 519-539.

Kolers, P. A. (1963). Interlingual word associations. *Journal of Verbal Learning & Verbal Behavior, 2,* 291-300.

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

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

小森早江子 (2015).「語彙連想課題における日本語と英語の相違について」. 中部大学 人文学部研究論集, 33, 69-79.

Lambert, W. E. (1956). Developmental aspects of second-language acquisition: I. Associational fluency, stimulus provocativeness, and word-order influence. *The Journal of Social Psychology, 43(1),* 83-89.

Laufer, B., & Aviad-Levitzky, T. (2017). What type of vocabulary knowledge predicts reading comprehension: Word meaning recall or word meaning recognition. *Modern Language Journal, 101(4),* 729-741.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language, 22,* 15-30.

Li, H. (2011). An investigation into the L2 mental lexicon of Chinese English learners by means of word association. *Chinese Journal of Applied Linguistics, 34(1),* 62-76.

Li, X., & Wang, W. (2016). Word class influence upon L1 and L2 English Word Association. *Chinese Journal of Applied Linguistics (De Gruyter), 39(4),* 440-458.

Loewen, S. (2015). *Introduction to Instructed Second Language Acquisition.* Routledge: New York.

Ma, Q., & Lee, H. Y. (2019). Measuring the Vocabulary Knowledge of Hong Kong Primary School Second Language Learners Through Word Associations: Implications for Reading Literacy. In: Reynolds B., Teng M. (eds) *English Literacy Instruction for Chinese Speakers.* Palgrave Macmillan, Singapore. https://doi.org/10.1007/978-981-13-6653-6_3

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88,* 375.

McDonald, K., & Asaba, M. (2015). "I don't know" use and guessing on the bilingual Japanese vocabulary size test: A preliminary report. *Vocabulary Learning and Instruction, 4,* 16-25.

McLean, S., Kramer, B., & Stewart, J. (2015). An empirical examination of the effect of guessing on vocabulary size test scores. *Vocabulary Learning and Instruction, 4,* 26-35.

McNamara, T. P., & Holbrook, J. B. (2003). *Semantic memory and priming.* In A. F. Healy & R. W. Proctor (Eds.), Handbook of psychology: Experimental psychology. New York: Wiley. Vol. 4, pp. 447-474.

Meyer, D. E., & Ruddy, M. G. (1974, April). Bilingual word recognition: Organization and retrieval of alternative lexical codes. Paper presented at the annual meeting of the Eastern Psychological Association, Philadelphia.

Meara, P. (1996). The classical research in L2 vocabulary acquisition. *Words, words, words: The translator and the language learner* (p. 38). Multilingual Matters: Clevedon.

Meara, P. (2009). *Connected words: Word associations and second language vocabulary acquisition.* Amsterdam, Netherlands: John Benjamins.

Nakayama, M., Ida, K., & Lupker, S. (2016). Cross-script L2-L1 noncognate translation priming in lexical decision depends on L2 proficiency: Evidence from Japanese–English bilinguals. *Bilingualism: Language and Cognition 19(5), 1001-1022.* DOI: 10.1017/S1366728915000462

Namei, S. (2004). Bilingual lexical development: A Persian–Swedish word association study. *International Journal of Applied Linguistics, 14(3),* 363-388.

Nation, I. S. P. (1990). *Teaching and learning vocabulary.* New York, NY: Newbury House.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge, UK: Cambridge University Press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review, 63,* 59-82.

Nation, I. S. P., & Beglar, D. (2007). *A vocabulary size test. The Language Teacher, 31,* 9-13.

Nation, I. S. P., & Webb, S. (2011). Content-based instruction and vocabulary learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning.* Volume II (pp. 631-644). New York: Routledge / Taylor & Francis.

Neely, J. H. (1991). *Semantic priming effects in visual word recognition: A selective review of current findings and theories.* In D. Besner & G. W. Humphreys (Eds.), Basic processes in reading: Visual word recognition. Hillsdale, NJ: Erlbaum. pp. 264-336.

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15,* 1003-1019.

Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is free association and what does it measure? *Memory & Cognition, 28,* 887-899.

Nissen, H. B. & Henriksen, B. (2006). Word class influence on word association test results. *International Journal of Applied Linguistics, 16,* 389-408.

Norrby, C. E., & Håkansson, G. (2007). Girl-lass or curl? Word associations in second language learners. *Australian Review of Applied Linguistics, 30(2),* 22.1-22.17.

Orita, M. (2002). Word associations of Japanese EFL learners and native speakers: Shifts in response type distribution and the associative development of individual words. *Review of English Language Education in Japan, 13,* 111-120.

Palermo, D. S. (1971). Characteristics of word association responses obtained from children in grades one through four. *Developmental Psychology, 5,* 118-123.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary development. In J. Coady & T. Huckin (Eds.),

Second language vocabulary acquisition: A rationale for pedagogy (pp. 174-200). Cambridge, UK: Cambridge University Press.

Petrey, S. (1977). Word associations and the development of lexical memory. *Cognition, 5,* 57-71.

Playfoot, D., Balint, T., Pandya, V., Parkes, A., Peters, M., & Richards, S. (2016). Are word association responses really the first words that come to mind? *Applied Linguistics, 39(5),* 607-624.

Playfoot, D., & Izura, C. (2013). Imageability, age of acquisition, and frequency factors in acronym comprehension. *Quarterly Journal of Experimental Psychology, 66(6),* 1131-1145.

Potter, M. C., So, K. F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning & Verbal Behavior, 23,* 23-38.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing, 10(3),* 355-371.

Read, J. (2000). *Assessing vocabulary.* Cambridge, UK: Cambridge University Press.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly, 10(1),* 77-89.

Saji, N., & Kajita, Y., & Imai, M. (2010). Learning verb system in second language: the influence of lexical knowledge of L1 in learning the relationships between the meanings of verbs in L2. *Second Language, 9,* 83-100.

Santos, A., Chaigneau, S. E., Simmons, W. K., & Barsalou, L. W. (2011). Property generation reflects word association and situated simulation. *Language and Cognition, 3(1),* 83-119.

Scarborough, D. L., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. *Journal of Verbal Learning & Verbal Behavior, 23,* 84-99.

Schmitt, N. (1998a). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning, 48(2),* 281-317.

Schmitt, N. (1998b). Quantifying word association responses: What is nativelike? *System, 26(3),* 389-401.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *Modern Language Journal, 95,* 26-43.

Schoonbaert, S., Duyck, W., Brysbaert, M., & Hartsuiker, R. J. (2009). Semantic and translation

priming from a first language to a second and back: Making sense of the findings. *Memory & Cognition, 37,* 569-86.

Shweta, Bajpai, R.C., & Chaturvedi, H.K. (2015). Evaluation of inter-rater agreement and inter-rater reliability for observational data: An overview of concepts and methods. *Journal of the Indian Academy of Applied Psychology, 41(3),* 20-27.

Smith, Y., Walters, J., & Prior, A. (2019). Target accessibility contributes to asymmetric priming in translation and cross-language semantic priming in unbalanced bilinguals. *Bilingualism: Language and Cognition, 22(1),* 157-176. doi: 10.1017/S1366728917000645

Sökmen, A. (1993). Word association results: A window to the lexicons of ESL students. *JALT Journal, 15(2),* 135-150.

Stewart, J. (2014). Do multiple-choice options inflate estimates of vocabulary size on the VST? *Language Assessment Quarterly, 11,* 271-282.

Stoeckel, T., Bennett, P., & McLean, S. (2016). Is "I don't know" a viable answer choice on the vocabulary size test? *TESOL Quarterly, 50,* 965-975.

Suzuki-Parker, J., & Higginbotham, G. (2019). Does method of administration influence word association test responses? *Vocabulary Education Research Bulletin, 8(1),* 11-16.

Van Hell, J. G., & de Groot, A. M. B. (1998a). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1(3),* 193-211.

Van Hell, J. G., & de Groot, A. M. B. (1998b). Disentangling Context Availability and Concreteness in Lexical Decision and Word Translation. *The Quarterly Journal of Experimental Psychology Section A, 51(1),* 41-63.

Wen, Y., & van Heuven, W. J. (2017). Non-cognate translation priming in masked priming lexical decision experiments: A meta-analysis. *Psychonomic Bulletin & Review, 24(3),* 879-886.

Wesche, M., & Paribakht, T. M. (1996). Assessing vocabulary knowledge: Depth vs. breadth. *Canadian Modern Language Review, 53,* 13-40.

White, L. (1989). *Universal grammar and second language acquisition* (pp.141-147). Philadelphia: John Benjamins Publishing Company.

Williams, J. N. (1994). The relationship between word meanings in the first and second language: Evidence for a common, but restricted, semantic code. *European Journal of Cognitive*

Psychology, 6, 195-220.

Wolter, B. (2001). Comparing the L1 and L2 mental lexicon: A depth of individual word knowledge model. *Studies in Second Language Acquisition, 23(1),* 41-69.

Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics, 32(4),* 430-449.

籔内智 (2015). 「自由連想課題を通して見た日本人英語学習者のメンタルレキシコン」. *京都精華大学紀要, 46,* 120-143.

Yokokawa, H., Yabuuchi, S., Kadota, S., Nakanishi, Y. & Noro, T. (2002). Lexical networks in L2 mental lexicon: Evidence from a word-association task for Japanese EFL learners. *Language Education & Technology, 39,* 21-39.

Zareva, A. (2007). Structure of the second language mental lexicon: How does it compare to native speakers' lexical organization? *Second Language Research, 23(2),* 123-153.

Zareva, A. (2011). Effects of lexical class and word frequency on L1 and L2 English-based lexical connections. *The Journal of Language Teaching and Learning, 1(2),* 1-17.

Zareva, A., Schwanenflugel, P. J., & Nikolova, Y. (2005). Relationship between lexical competence and language proficiency: Variable sensitivity. *Studies in Second Language Acquisition, 27(4),* 567-595.

Zareva, A., & Wolter, B. (2012). The 'promise' of three methods of word association analysis to L2 lexical research. *Second Language Research, 28(1),* 41-67.

Zhang, D., & Koda, K. (2017). Assessing L2 vocabulary depth with word associates format tests: Issues, findings, and suggestions. *Asian-Pacific Journal of Second and Foreign Language Education, 2(1),* 1-30.

Zhang, X. (2013). The I don't know option in the vocabulary size test. *TESOL Quarterly, 47,* 790-811.

ACKNOWLEDGMENTS

I would like to express my sincere thanks to all people who have helped and encouraged me during my research and completion of this doctoral dissertation.

First and foremost, I would like to express my deep gratitude to my supervisor Professor Yokokawa Hirokazu for his helpful guidance and warm encouragement. I have been very blessed to have many informative suggestions from him which helped me shape my research directions and approaches and achieve the goal.

I would also like to express my appreciation to all members in our laboratory who gave me lots of useful advice about the study in our seminar. Special thanks go to Xie Zhanmei, my senior lab mate, who talked a lot with me about not only academic but also life, and always encouraged me.

The completion of this dissertation was also supported by my beloved teacher in undergraduate period Professor Wang Jiming, my cousin Jin Shuran, my friends Xu Min, Wang Yifang, Li Mingxiao, Li Yifei, and so on. They helped me a lot with data collection. As I cannot return China to do all for the corona virus pandemic, they helped me contact participants, or scan the collected data of participants to make sure that I could receive it in time and securely.

Finally, I hope to express my most profound gratitude to my family for their understanding, encouragement, and all-around support.

Although it is hard to list all the names who have helped me, I am fortunate to have all the help and owe a debt of gratitude to all of them. This work could never have been completed without the help from them all.

APPENDICES

★★★

Appendix A Word Association Task & English Learning Background Questionnaire (Version 1)

关于词汇联想、词汇网的调查

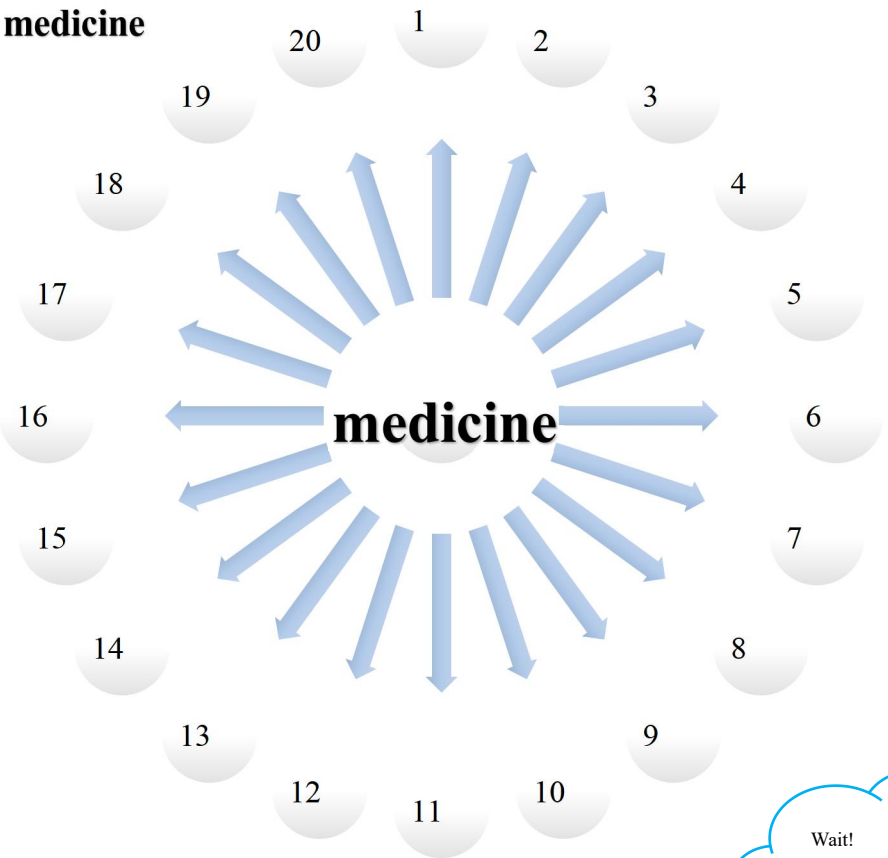
同学你好！感谢配合。
本次调查包括词汇联想（英语）、关于英语学习经历等相关内容的问卷、词汇联想（中文）三大部分。
●请注意：
在回答问题的过程中，需要自主计时。请准备开启手机的秒表功能，根据之后的文字提示，完成计时操作。**请不要使用包括查词等在内的手机的其他功能。**
感谢配合。

下面正式开始。

一、词汇联想（英语）

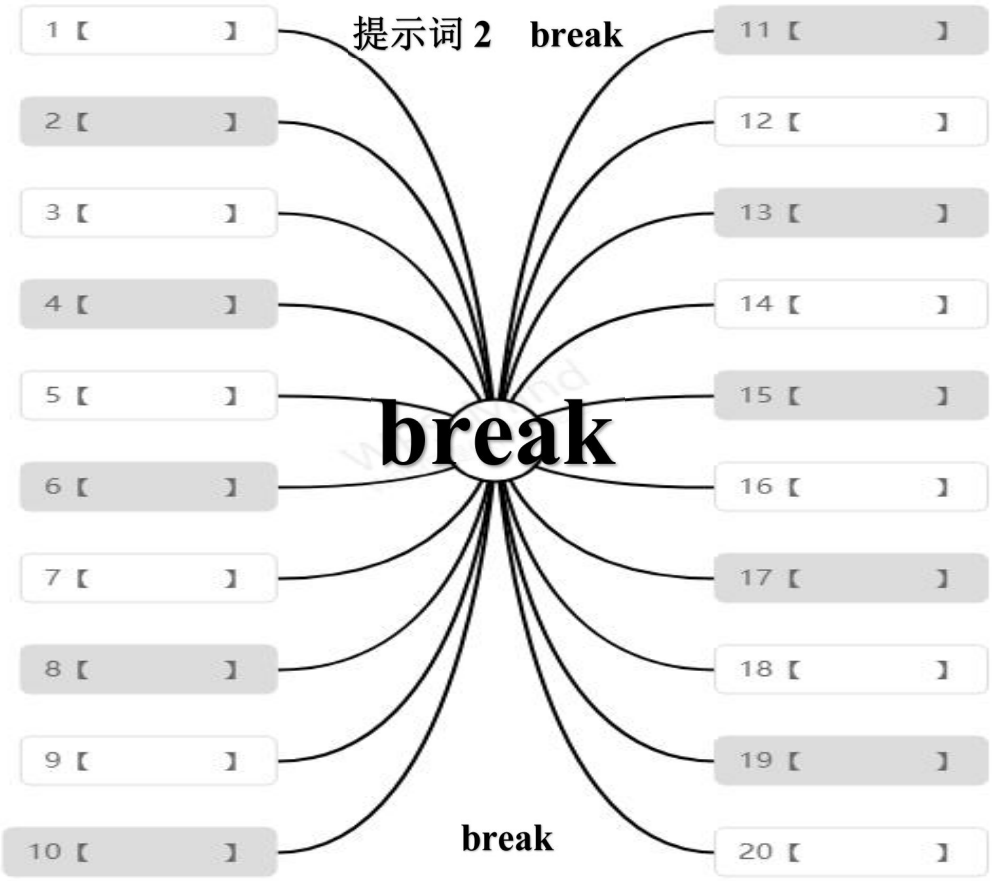
此调查的目的是了解脑内的英语词汇网的构造。请填写由提示单词所联想到的英文单词。本题共有 **9 个**提示词，每词联想时间为 **2 分钟**。请自主计时，分别在 2 分钟内**尽可能多**地联想相关单词并填入提示词周边的空白处。请从 1 号空白处开始依次填写。
例：若提示词为 apple、你可能联想到的单词有：red, fruit, orange, tree, sweet, eat，iPhone...

提示词 1　medicine

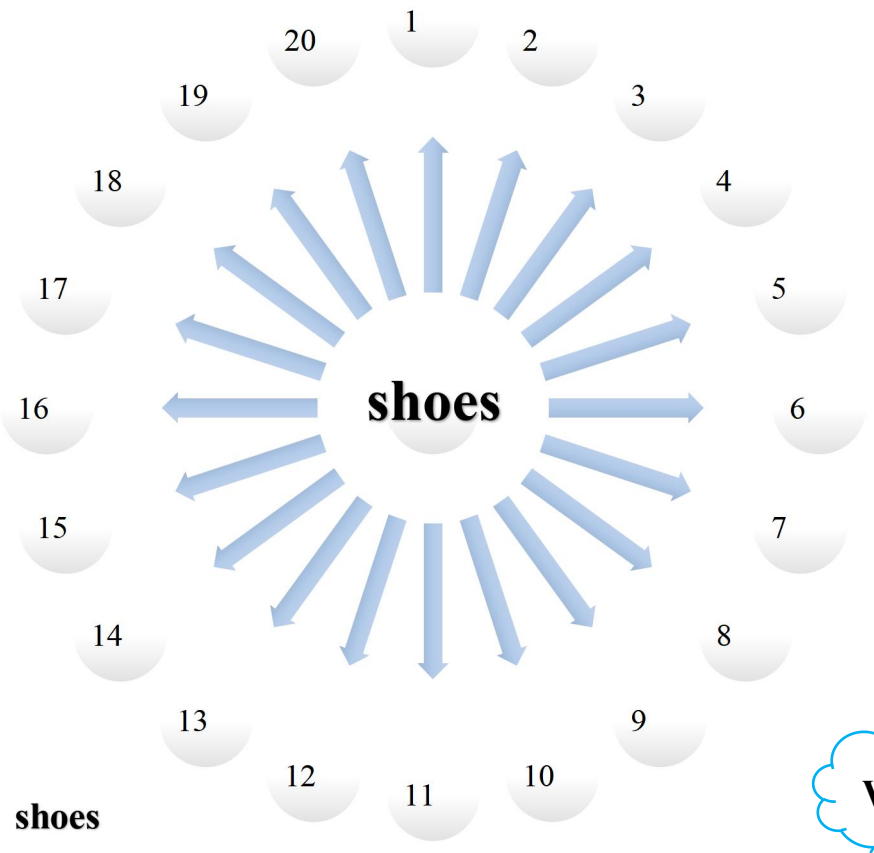


Wait!
请等待 2 分钟时间到后再开始下一个单词的联想

★★★



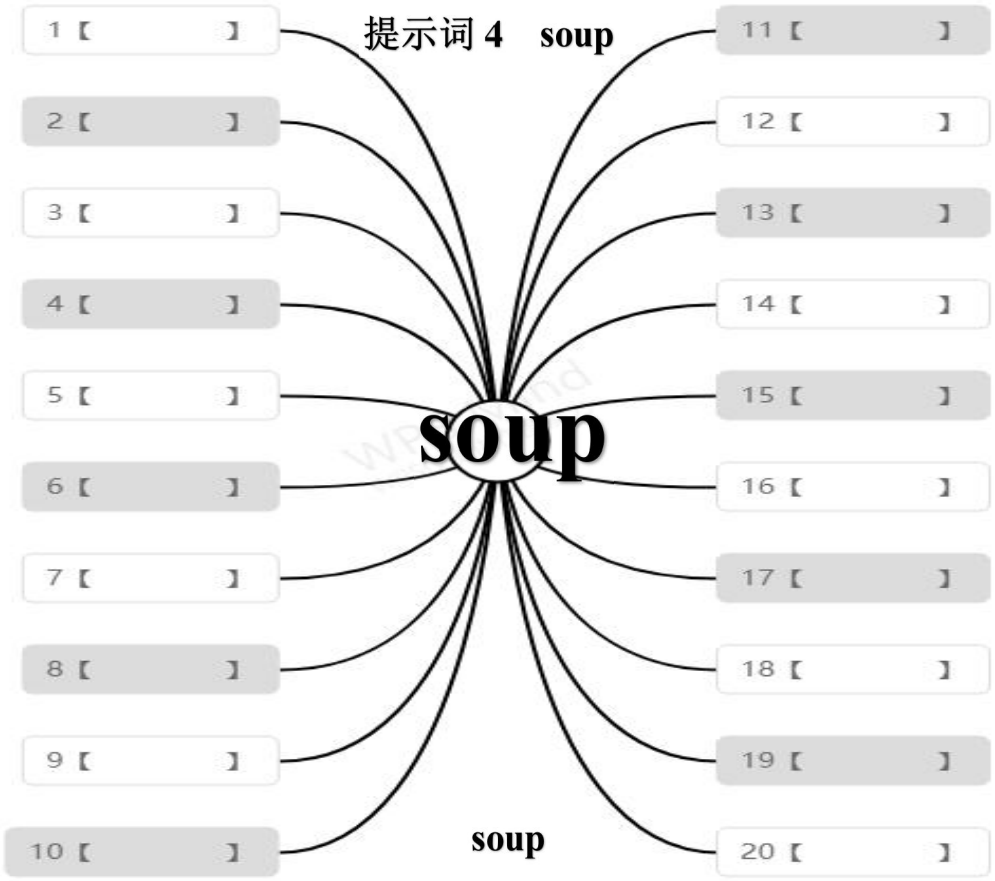
Wait!



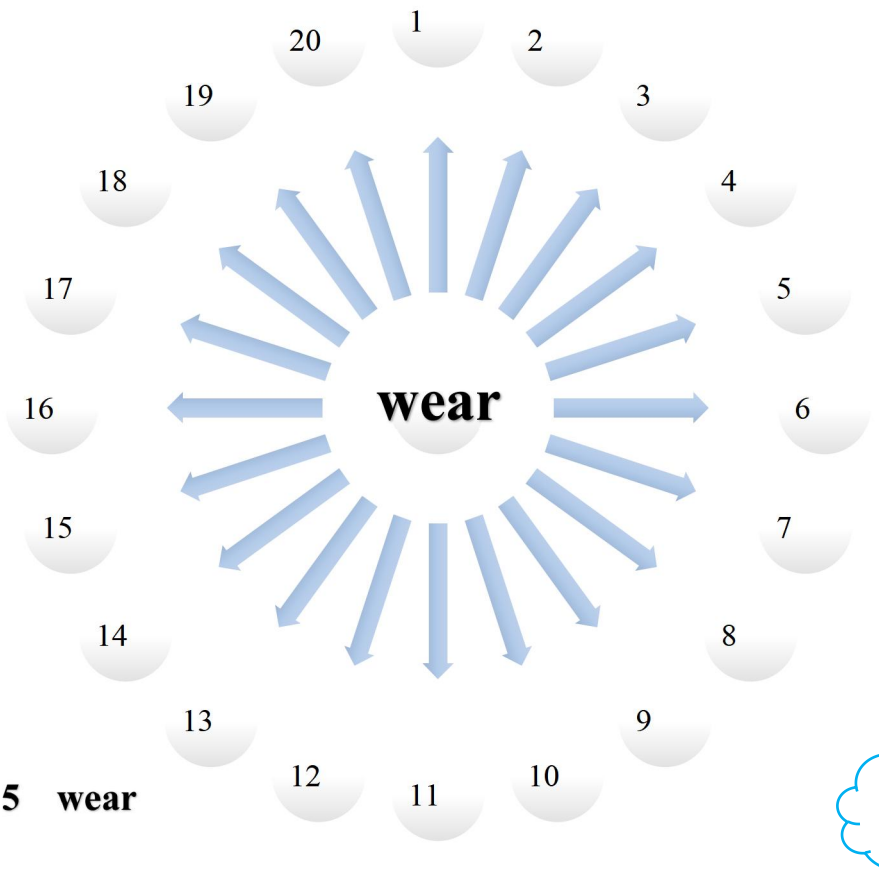
提示词 3　shoes

Wait!

★★★



Wait!

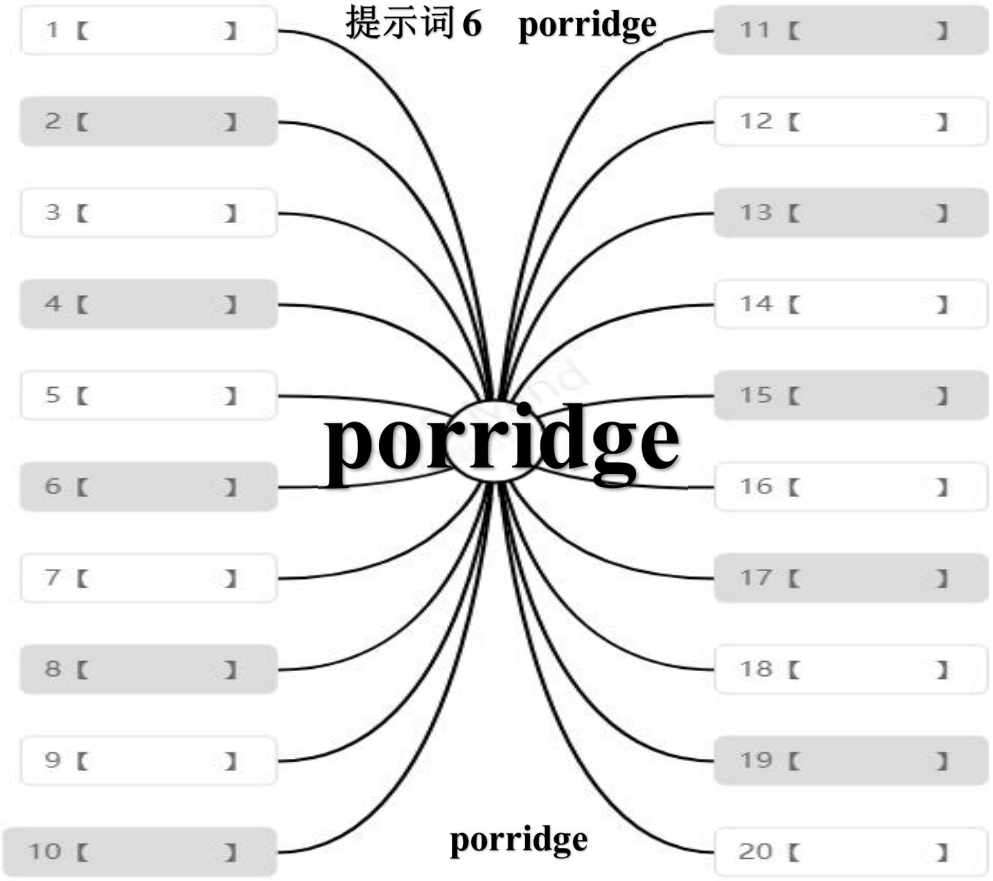


提示词 5　wear

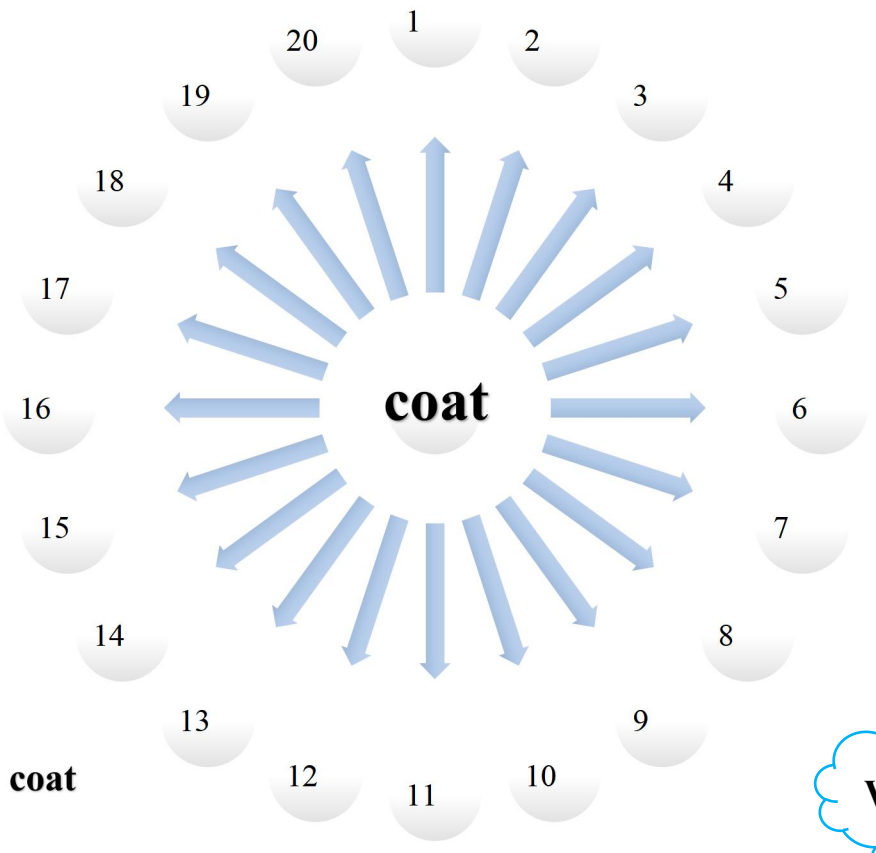
Wait!

★★★

提示词 6　porridge



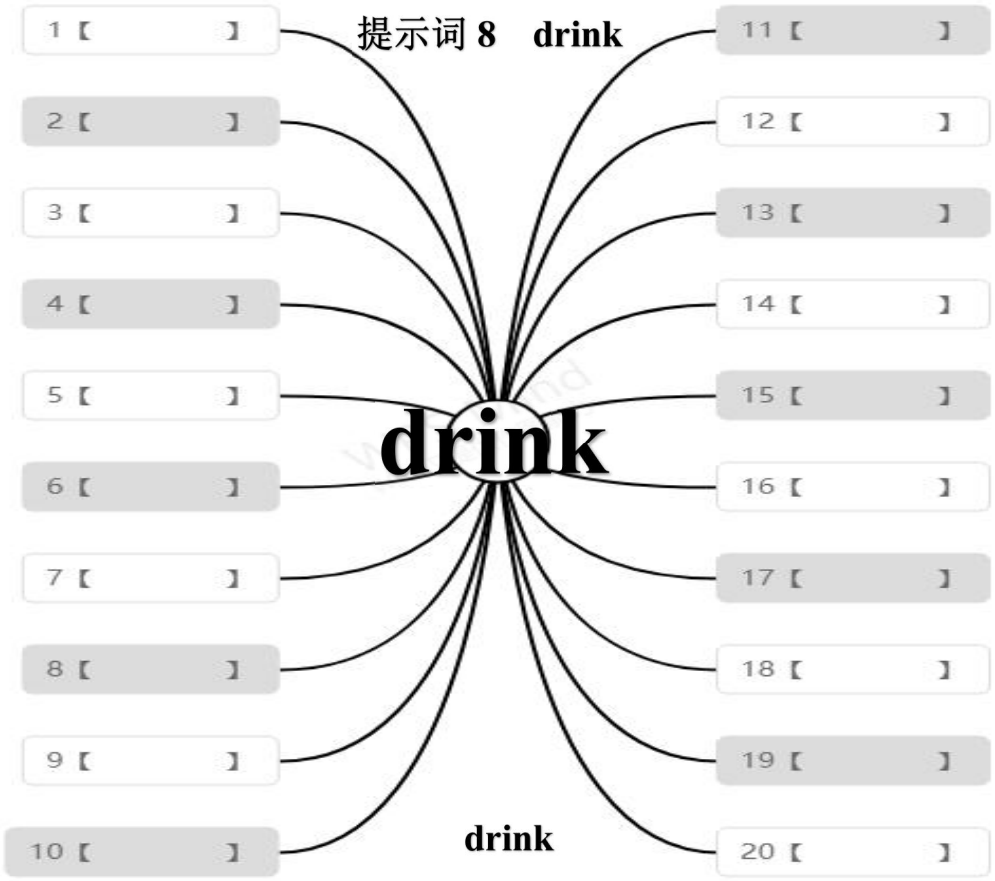
Wait!



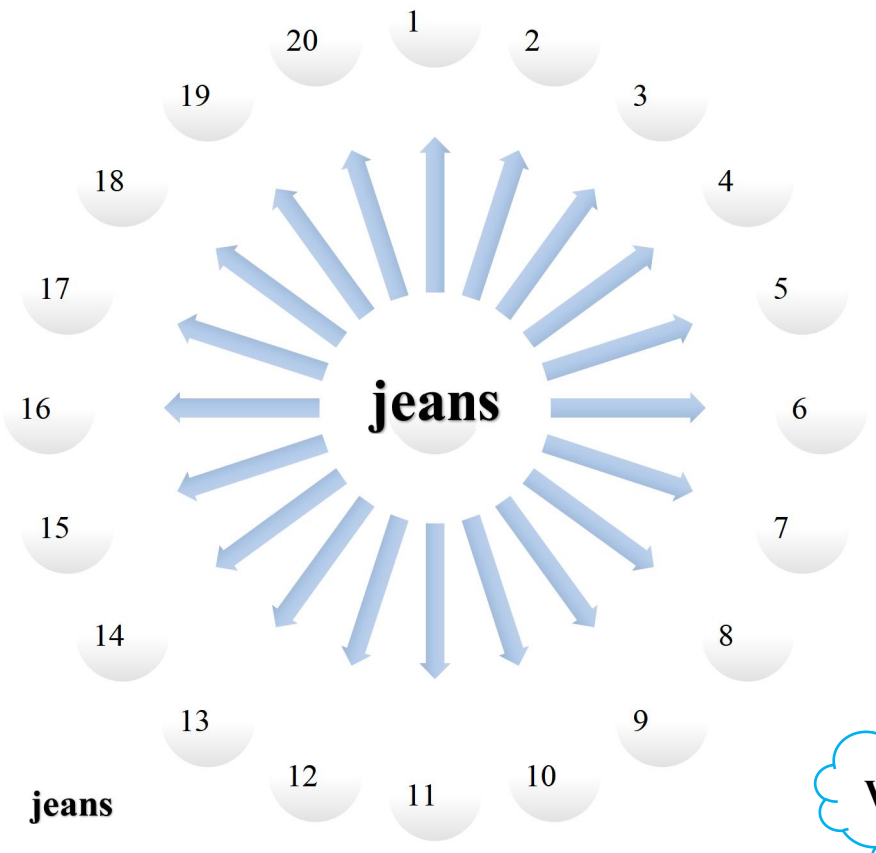
提示词 7　coat

Wait!

★★★


提示词 8 drink
1 【 】
2 【 】
3 【 】
4 【 】
5 【 】
6 【 】
7 【 】
8 【 】
9 【 】
10 【 】
11 【 】
12 【 】
13 【 】
14 【 】
15 【 】
16 【 】
17 【 】
18 【 】
19 【 】
20 【 】
drink
drink


Wait!


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
jeans

提示词 9 jeans


Wait!

★★★

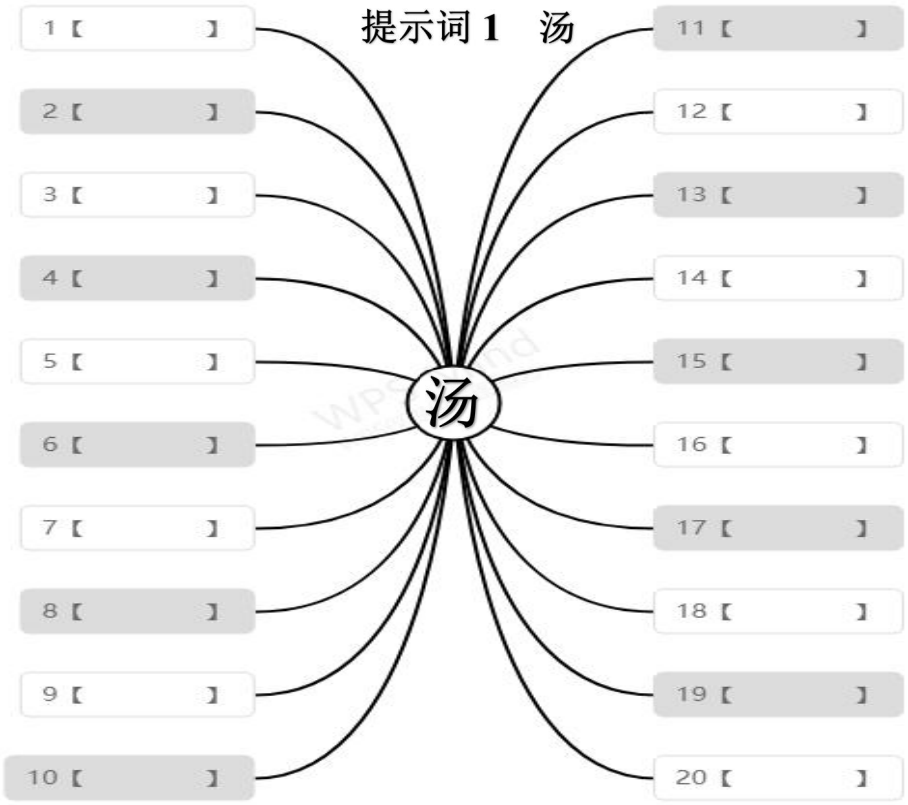
二、关于英语学习经历等相关内容的问卷

	2020 年　12 月　___日　_____专业　____年级
	下面询问有关你的英语学习经历等内容的问题。请填写划线处或者在最恰当的数字编号处画圈。
1	性别　1.男性　2.女性
2	年龄 _____岁
3	英语学习年数　____年____个月
4	英语学习开始时间（例：小学三年级开始）　_____
5	是否有英语以外的外语学习经历？　1.有　2.没有
6	请在 5 中选择了 1.有英语以外的外语学习经历的同学回答。 请填写所学语言_____语，及学习时间：____年____个月
7	是否有过出国经历？　1.有　2.没有
8	请在 7 中选择了 1.有过出国经历的同学回答。请填写相关经历。 去过哪个国家？去了多长时间？ △注意：请填写所有的出国经历　填入例：美国▪一年零三个月 _____　_____ _____　_____
9	下面询问英语考试的分数。 请填写参加过的考试的成绩和取得相应成绩的时间。未参加者请在未参加处画圈。 CET4 大学英语四级　_____分　_____年_____月　未参加 CET6 大学英语六级　_____分　_____年_____月　未参加 IELTS 雅思　_____分　_____年_____月　未参加 TOEFL 托福　_____分　_____年_____月　未参加 外国语学院的同学请填写以下内容： TEM4 英语专业四级　_____分　_____年_____月　未参加 TEM8 英语专业八级　_____分　_____年_____月　未参加
10	此次调查结束后，根据需要可能会对您所填写的内容进行回访确认。 如果愿意协助完成后续调查，请填写姓名及联系方式（邮箱）。 姓名　邮箱

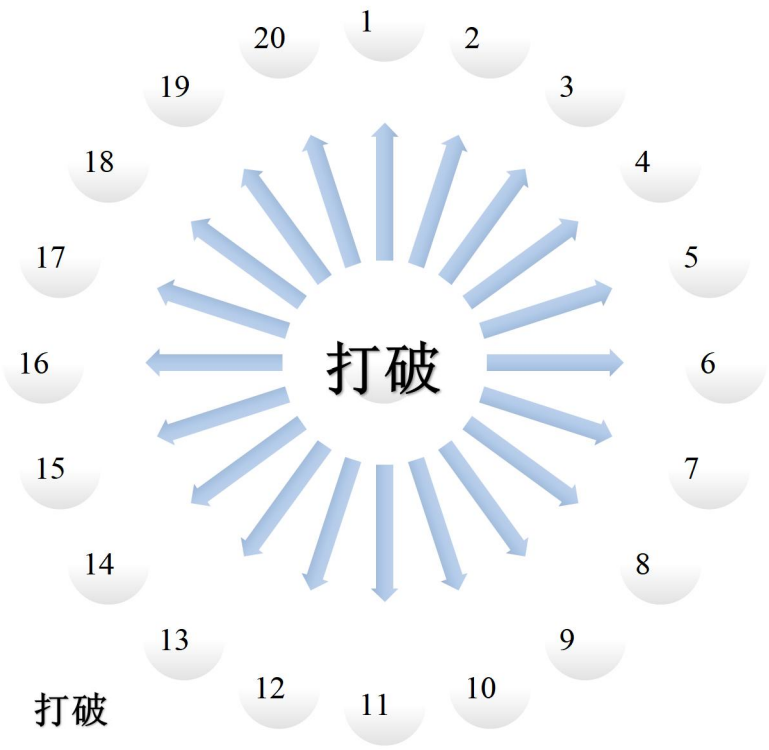
★★★

三、词汇联想（中文）

此调查的目的是了解脑内的汉语词汇网的构造。请写出由提示词所联想到的汉语词语。本题共有 **10 个**提示词，每词联想时间为 **2 分钟**。请自主计时，分别在 2 分钟内**尽可能多**地联想相关单词并填入空白处。请从 1 号空白处开始依次填写。



Wait!
请等待 2 分钟到的指示后再开始下一个单词的联想

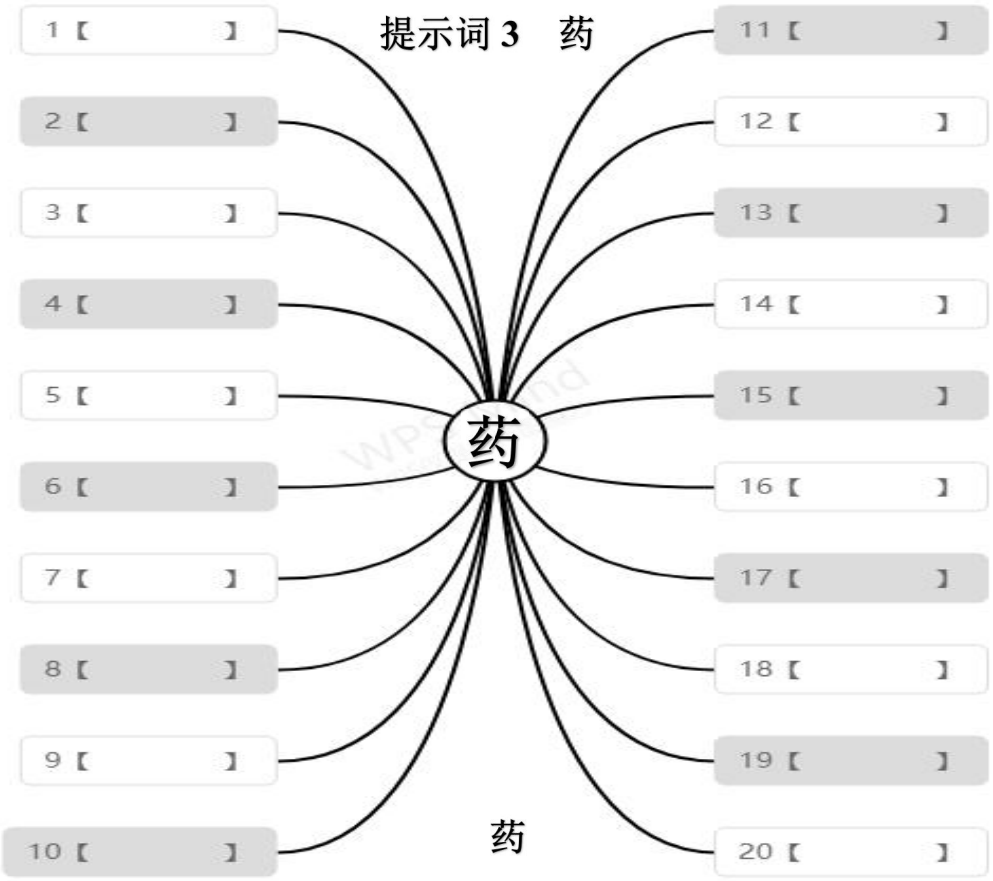


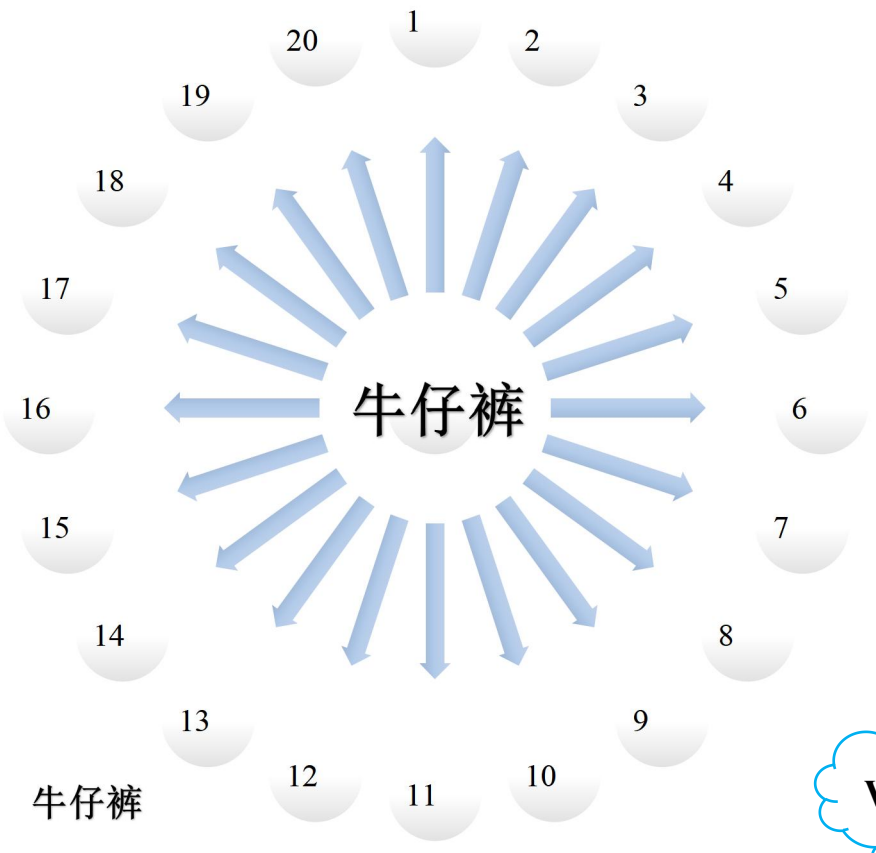
提示词 2　打破

Wait!

★★★

提示词 3　药

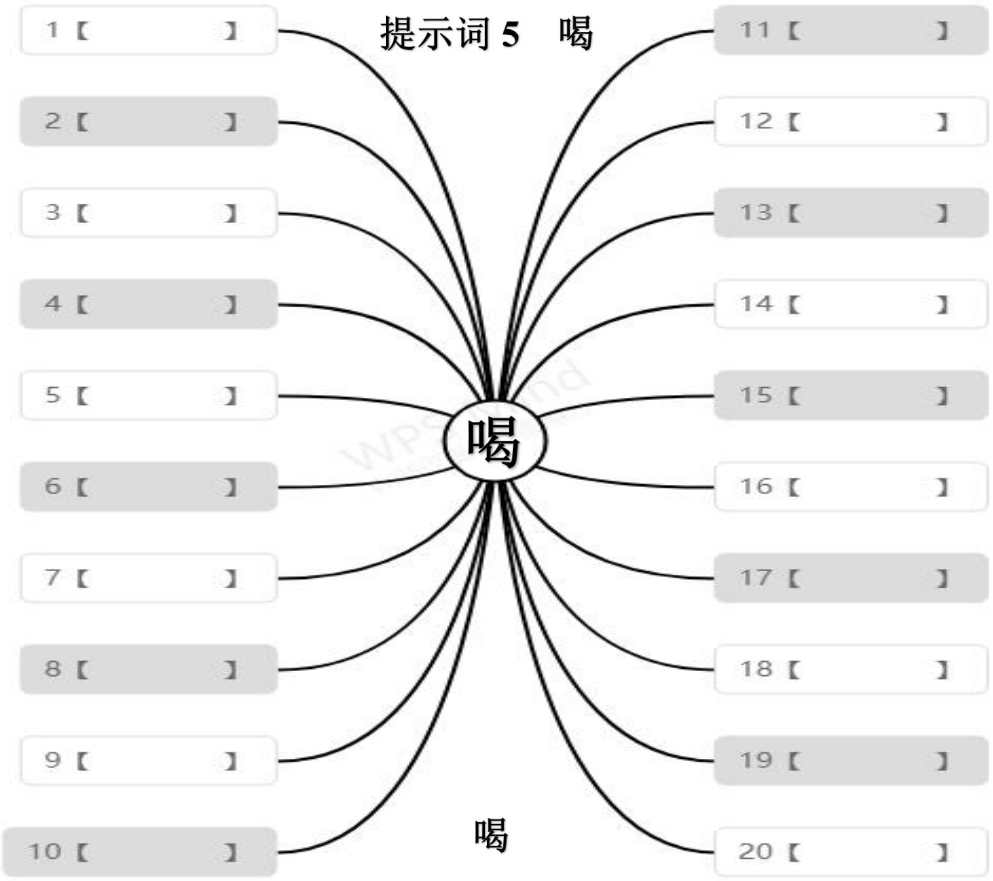




提示词 4　牛仔裤

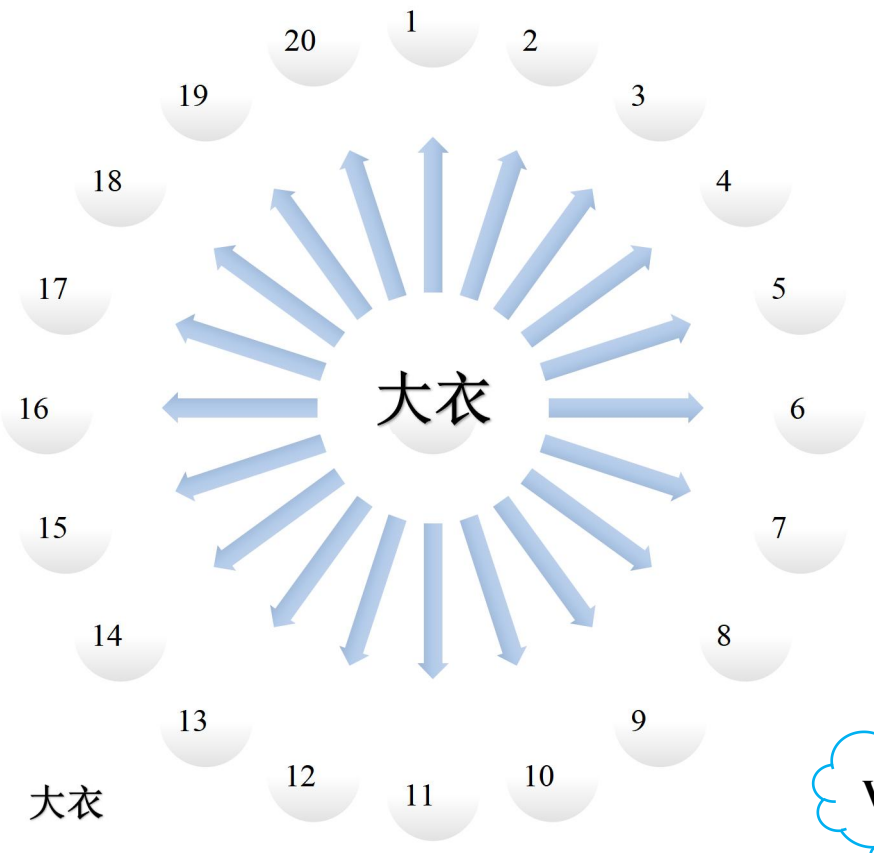
★★★

提示词 5　喝



Wait!

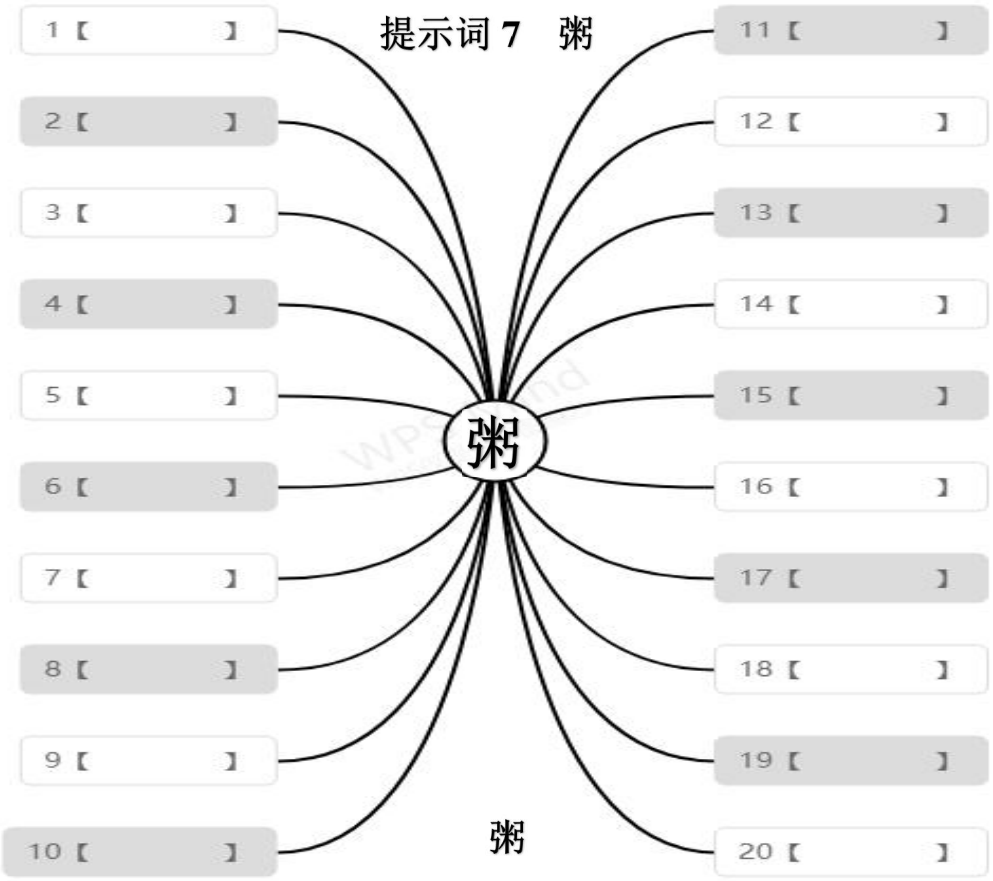
提示词 6　大衣



Wait!

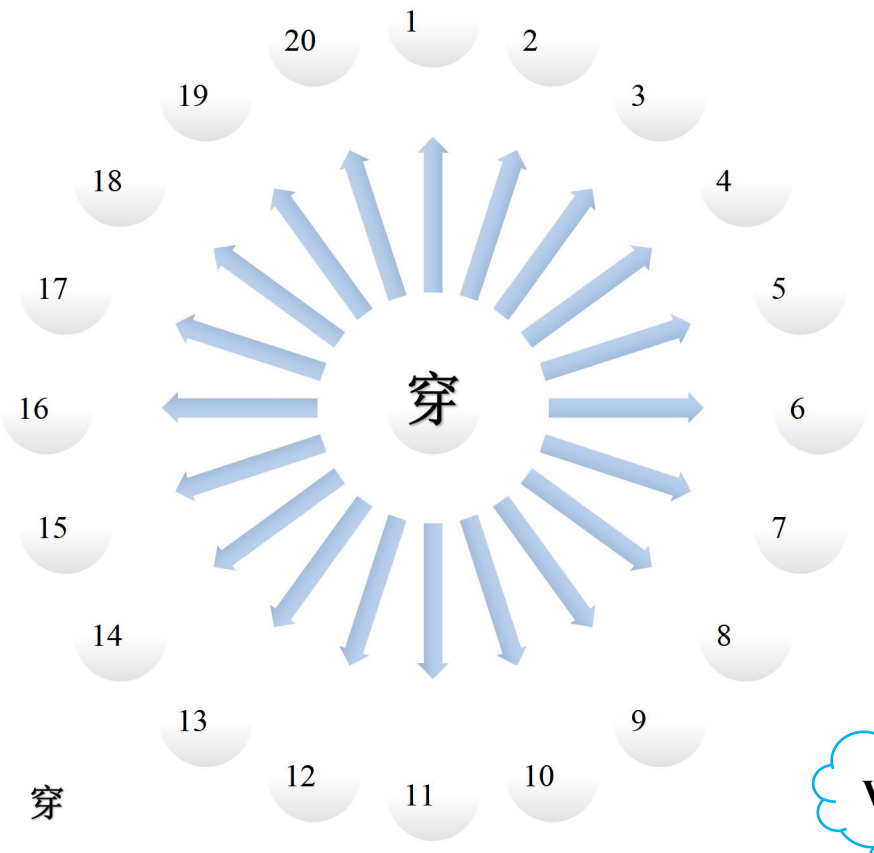
★★★

提示词 7　粥



粥

Wait!

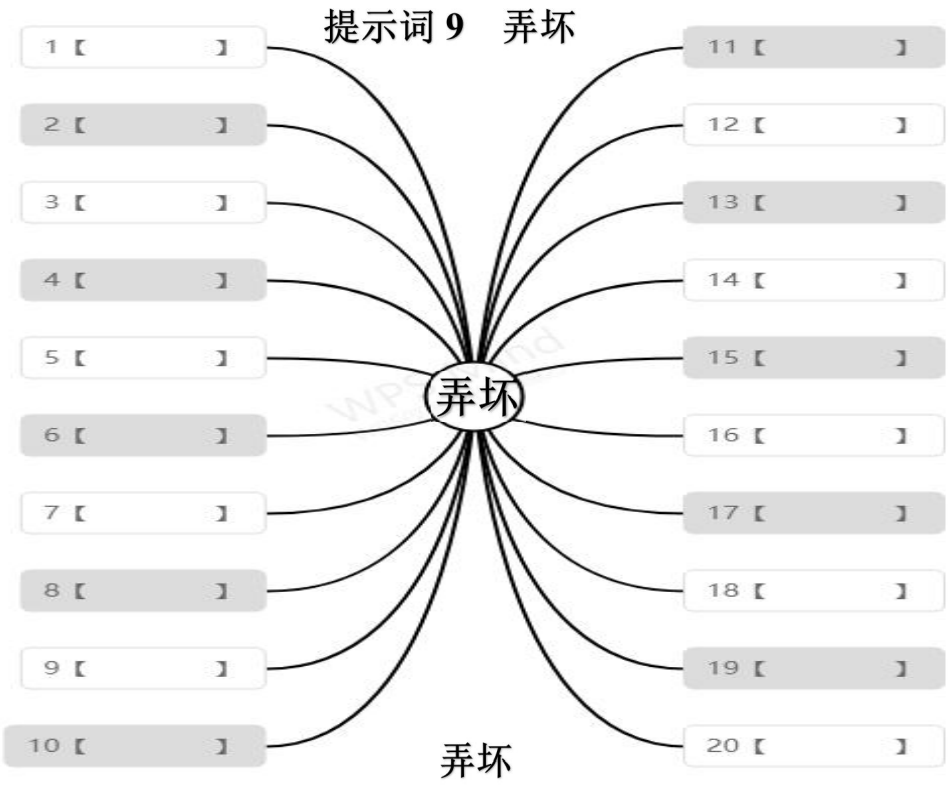


提示词 8　穿

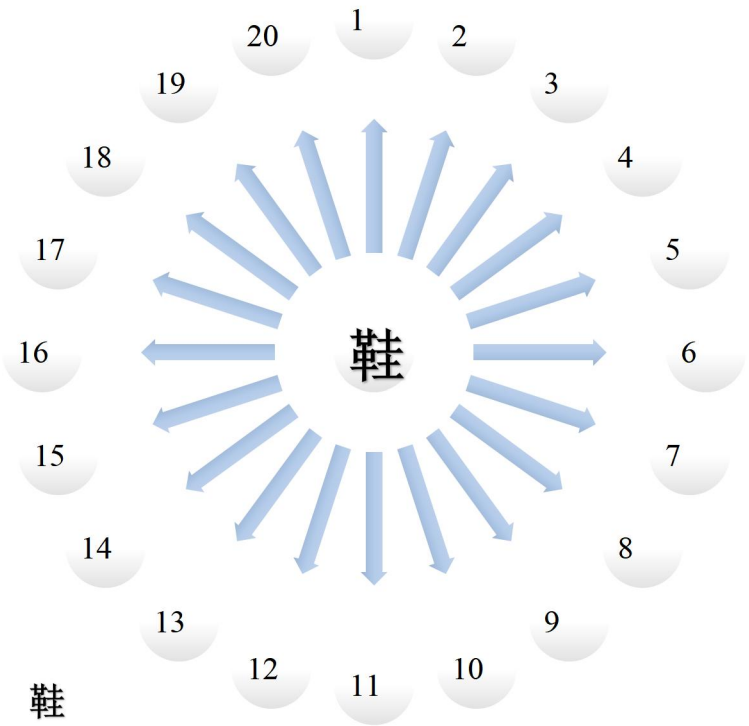
Wait!

★★★

提示词 9　弄坏



提示词 10　鞋



以上是此次调查的全部内容。再次感谢您的配合。

关于此次调查，如果您有任何疑问等，请联系我：
神户大学国际文化学研究科　毕奇　　邮箱：bi@stu.kobe-u.ac.jp
衷心感谢！谢谢！祝生活愉快，学习进步。

Appendix B Written Consent for Experiment 1

同意书

本调查作为博士研究的一环展开。研究目的是通过对英语学习者的脑内词汇网的研究，探讨其英语词汇学习的特点。

调查研究实施者会妥善保管所有调查数据，充分保护参与调查者的隐私。在最终的研究成果展示等公开环节中，也会充分注意调查参与者的姓名等相关隐私不被泄露。研究者会在调查数据的基础上完成论文并提交给所属神户大学。研究成果将会在相关学术研讨会上发表。调查数据不会用于本研究以外的其他目的。

如果有幸能够得到您的配合，请在下方签字。在此深表谢意。

参加者（签字） _____

日期：2020 年 12 月　　日

您如果有关于此调查的任何问题，请通过以下联系方式与我取得联系。

研究实施者:神户大学国际文化学研究科　毕奇

邮箱:bi@stu.kobe-u.ac.jp

Appendix C VST

词汇量测试

First 1000

1. see: They **saw** it.
   a. 切
   b. 等待
   c. 看
   d. 开始

2. time: They have a lot of**time**.
   a. 钱
   b. 食物
   c. 时间
   d. 朋友

3. period: It was a difficult **period**.
   a. 问题
   b. 时间
   c. 要做的事情
   d. 书

4. figure: Is this the right **figure**?
   a. 答案
   b. 地方
   c. 时间
   d. 号码

5. poor: We **are poor**.
   a. 贫穷的
   b. 感到幸福的
   c. 很感兴趣的
   d. 不喜欢努力工作的

6. drive: He **drives** fast.
   a. 游泳
   b. 学习
   c. 扔球
   d. 开车

7. jump: She tried to **jump**.
   a. 漂浮
   b. 跳
   c. 停车
   d. 跑

8. shoe: Where is your **shoe**?
   a. 父或母
   b. 钱包
   c. 钢笔
   d. 鞋子

9. standard: Her **standards** are very high.
   a. 后跟
   b. 分数
   c. 要价
   d. 标准

10. basis: I don't understand the **basis**.
    a. 原因
    b. 话
    c. 路标
    d. 中心议题

Second 1000

1. maintain: Can they **maintain** it?
   a. 维持
   b. 扩大
   c. 改善
   d. 得到

2. stone: He sat on a **stone**.
   a. 石头
   b. 凳子
   c. 垫子
   d. 树枝

3. upset: I am **upset**.
   a. 疲倦的
   b. 著名的
   c. 富足的
   d. 不高兴的

4. drawer: The **drawer** was empty.
   a. 抽屉
   b. 车库
   c. 冰箱
   d. 鸟笼

5. patience: He **has no patience**.
   a. 没有耐心
   b. 很忙
   c. 没有信心
   d. 不公正

6. nil: His mark for that question was **nil**.
   a. 很差的
   b. 什么也没有的
   c. 很好的
   d. 中等的

7. pub: They went to the **pub**.
   a. 酒吧
   b. 银行
   c. 商场
   d. 游泳池

8. circle: Make a **circle**.
   a. 素描
   b. 空白
   c. 圆圈
   d. 大洞

9. microphone: Please use the **microphone**.
   a. 微波炉
   b. 麦克风
   c. 显微镜
   d. 手机

10. pro: He's a **pro**.
    a. 间谍
    b. 傻瓜
    c. 记者
    d. 职业运动员

Third 1000

1. soldier: He is a **soldier**.
   a. 商人
   b. 学生
   c. 金属工艺制造者
   d. 士兵

2. restore: It has been **restored**.
   a. 重复
   b. 重新分配
   c. 降价
   d. 复原

3. jug: He was holding a **jug**.
   a. 罐子
   b. 聊天
   c. 贝雷帽
   d. 枪

4. scrub: He is **scrubbing** it.
   a. 抓
   b. 修理
   c. 刷洗
   d. 作素描

5. dinosaur: The children were pretending to be **dinosaurs**.
   a. 海盗
   b. 仙女
   c. 龙
   d. 恐龙

6. strap: He broke the **strap**.
   a. 诺言
   b. 盖子
   c. 盘子
   d. 皮带

7. pave: It was **paved**.
   a. 堵塞
   b. 分开
   c. 镶金边
   d. 铺路

8. dash: They **dashed** over it.
   a. 猛冲
   b. 磨蹭
   c. 争吵
   d. 瞥见

9. rove: He couldn't stop **roving**.
   a. 喝醉
   b. 漂泊
   c. 哼曲子
   d. 努力工作

10 lonesome: He felt **lonesome**.
a. 不领情的
b. 疲倦的
c. 孤独的
d. 精力充沛的

Fourth 1000

1. compound: They made a new **compound**.
a. 协议
b. 复合物
c. 公司
d. 预言

2. latter: I agree with the **latter**.
a. 牧师
b. 理由
c. 后者
d. 答案

3. candid: Please be **candid**.
a. 小心的
b. 表示同情的
c. 公平的
d. 直率的

4. tummy: Look at my **tummy**.
a. 围巾
b. 肚子
c. 松鼠
d. 拇指

5. quiz: We made a **quiz**.
a. 箭筒
b. 错误
c. 测验
d. 鸟巢

6. input: We need more **input**.
a. 输入
b. 工人
c. 填料
d. 钱

7. crab: Do you like **crabs**?
a. 蟹
b. 薄脆饼干
c. 又紧又硬的领子
d. 蟋蟀

8. vocabulary: You will need more **vocabulary**.
a. 词汇
b. 技巧
c. 钱
d. 枪

9. remedy: We found a good **remedy**.
a. 矫正问题的方法
b. 餐馆
c. 食谱
d. 等式

10. allege: They **alleged it**.
a. 辩解
b. 剽窃
c. 证明
d. 反抗

Fifth 1000

1. deficit: The company **had a large deficit**.
a. 出现赤字
b. 贬值
c. 有这笔大开销的计划
d. 在银行里有很多存款

2. weep: He **wept**.
a. 毕业
b. 哭
c. 死
d. 担心

3. nun: We saw a **nun**.
a. 蠕虫
b. 事故
c. 修女
d. 天空中无法解释的亮光

4. haunt: The house is **haunted**.
   a. 充满了装饰物
   b. 已被出租
   c. 空的
   d. 闹鬼

5. compost: We need some **compost**.
   a. 大力支持
   b. 扶持
   c. 混凝土
   d. 堆肥

6. cube: I need one more **cube**.
   a. 大头针
   b. 立方体
   c. 缸子
   d. 卡片

7. miniature: It is a **miniature**.
   a. 微型画
   b. 显微镜
   c. 微生物
   d. 在书法中把字母连在一起细小的连线

8. peel: Shall I **peel** it?
   a. 浸泡
   b. 削皮
   c. 烫洗
   d. 切成薄片

9. fracture: They found a **fracture**.
   a. 裂口
   b. 碎片
   c. 夹克衫
   d. 稀有的宝石

10. bacterium: They didn't find a single **bacterium**.
    a. 细菌
    b. 开有红色或橘黄色花的植物
    c. 骆驼
    d. 赃物

Sixth 1000

1. devious: Your plans are **devious**.
   a. 诡计多端的
   b. 成熟的
   c. 考虑不周详的
   d. 过于昂贵的

2. premier: The **premier** spoke for an hour.
   a. 律师
   b. 讲师
   c. 冒险家
   d. 总理

3. butler: They have a **butler**.
   a. 男管家
   b. 锯
   c. 家庭教师
   d. 地窖

4. accessory: They gave us some **accessories**.
   a. 签证
   b. 官方命令
   c. 选择
   d. 额外的零件

5. threshold: They raised the **threshold**.
   a. 旗子
   b. 门槛
   c. 天花板
   d. 利息

6. thesis: She has completed her **thesis**.
   a. 论文
   b. 归纳证词
   c. 试用期
   d. 延期治疗

7. strangle: He **strangled** her.
   a. 掐死
   b. 宠坏
   c. 绑架
   d. 赞美

8. cavalier: He treated her in a **cavalier** manner.
   a. 慢待的
   b. 礼貌的

c. 尴尬的
d. 兄长的

9. malign: His **malign** influence is still felt.
a. 邪恶的
b. 好的
c. 非常重要的
d. 秘密的

10. veer: The car **veered**.
a. 改变方向或路线
b. 剧烈晃动
c. 发生逆火引起爆鸣
d. 打滑

Seventh 1000

1. olive: We bought **olives**.
a. 橄榄
b. 康乃馨
c. 男人的游泳衣
d. 清除杂草的工具

2. quilt: They made a **quilt**.
a. 遗嘱
b. 合同
c. 被子
d. 羽毛笔

3. stealth: They did it by **stealth**.
a. 花费大量的钱
b. 逼迫
c. 悄悄的或秘密的行动
d. 没有注意到所遇到的问题

4. shudder: The boy **shuddered**.
a. 低语
b. 差点摔倒
c. 发抖
d. 大声叫喊

5. bristle: The **bristles** are too hard.
a. 问题
b. 短而硬的毛发
c. 折叠床
d. 鞋底

6. bloc: They have joined this **bloc**.
a. 乐队
b. 小偷帮
c. 侦察员
d. 集团

7. demography: This book is about **demography**.
a. 土地使用模式研究
b. 用图片表示数字事实的研究
c. 水文学
d. 人口学

8. gimmick: That's a good **gimmick**.
a. 高空作业时所站的东西
b. 钱包
c. 引人注意的行为或事物
d. 花招

9. azalea: This **azalea** is very pretty.
a. 杜鹃花
b. 由天然棉所制成的很轻的材料
c. 莎丽
d. 扇贝

10. yoghurt: This **yoghurt** is disgusting.
a. 淤泥
b. 伤口
c. 酸奶
d. 榲桲

Eighth 1000

1. erratic: He was **erratic**.
a. 完美的
b. 很坏的
c. 很有礼貌的
d. 不可靠的

2. palette: He lost his **palette**.
a. 装鱼的篮子
b. 胃口
c. 年轻的女伴
d. 调色板

3. null: His influence **was null**.
   a. 具有好的结果
   b. 毫无帮助的
   c. 没有效果的
   d. 持久的

4. kindergarten: This is a good **kindergarten**.
   a. 消遣
   b. 幼儿园
   c. 背包
   d. 图书馆

5. eclipse: There was an **eclipse**.
   a. 飓风
   b. 泼溅
   c. 大屠杀
   d. 日食

6. marrow: This is the **marrow**.
   a. 吉祥物
   b. 骨髓
   c. 操纵杆
   d. 增加工资

7. locust: There were hundreds of**locusts**.
   a. 飞蝗
   b. 志愿者
   c. 素食者
   d. 颜色鲜艳的野花

8. authentic: It is **authentic**.
   a. 真的
   b. 非常吵闹的
   c. 老的
   d. 干旱的

9. cabaret: We saw the **cabaret**.
   a. 壁画
   b. 歌舞表演卡巴莱
   c. 蟑螂
   d. 美人鱼

10. mumble: He started to **mumble**.
    a. 集中精力
    b. 颤抖
    c. 远远地落后于其他人
    d. 咕哝

Ninth 1000

1. hallmark: Does it have a **hallmark**?
   a. 表明什么时候之前应该被使用的印记
   b. 纯度印记
   c. 表明经过皇室同意的标记
   d. 禁止复制的标记

2. puritan: He is a **puritan**.
   a. 喜欢被人注意的人
   b. 具有严格道德标准的人
   c. 吉普赛人
   d. 守财奴

3. monologue: Now he has a **monologue**.
   a. 单眼镜
   b. 独白
   c. 专制
   d. 把字母有趣地连在一起的图画

4. weir: We looked at the **weir**.
   a. 行为古怪的人
   b. 红树属植物
   c. 通过吹来演奏的古老金属乐器
   d. 拦河坝

5. whim: He had lots of**whims**.
   a. 古老的金币
   b. 母马
   c. 没有任何动机的奇异想法
   d. 疼痛的红肿块

6. perturb: I was **perturbed**.
   a. 被迫同意的
   b. 烦恼的
   c. 困惑的
   d. 湿透的

7. regent: They chose a **regent**.
   a. 不负责人的人
   b. 暂时主持会议的人
   c. 摄政者
   d. 代表

8. octopus: They saw an **octopus**.
a. 猫头鹰
b. 潜水艇
c. 直升飞机
d. 章鱼

9. fen: The story is set in the **fens**.
a. 沼泽
b. 山地
c. 贫民窟
d. 很久以前

10. lintel: He painted the **lintel**.
a. 过梁，楣
b. 渡船
c. 长有伸展树枝和绿色果实的美丽的树
d. 戏院中显示场景的板子

Appendix D Consent for Experiment 2 Online

同意书

您好，非常感谢您的参与！

本调查作为博士论文研究的一环展开。研究目的是通过对词汇知识和脑内词汇网的研究，探讨中国英语学习者的英语词汇学习特点。

调查研究的实施者会妥善保管所有调查数据，充分保护参与调查者的隐私。在最终的研究成果展示等公开环节中，也会充分注意参与者的姓名等相关隐私不被泄露。调查数据会被统一整理分析，不存在针对个人的英语水平等的评判。

研究者会在所收集调查数据的基础上完成论文并提交给所属神户大学。研究成果将会在相关学术研讨会上发表。调查数据不会用于本研究以外的其他目的。

如果有幸能够得到您的配合，请在“同意署名”处输入姓名，然后进入调查部分。

本调查包括词汇联想、词汇知识深度测试、词汇量测试以及关于英语学习经历等相关内容的问卷四大部分。

同意署名 （输入姓名） _____

关于此次调查，如果您有其他任何疑问等，请联系我：

研究实施人：神户大学国际文化学研究科 毕奇

邮箱：bi@stu.kobe-u.ac.jp

Appendix E The Word List of Collocation Decision Task

	L2	L2	L2	L2		L1	L1	L1	L1	
L1>L2	drink	water	eat	water	*	喝	水	吃	水	*
	eat	soup	drink	soup	*	喝	汤	吃	汤	*
	take	medicine	eat	medicine	*	吃	药			
			drink	medicine	*	喝	药			
	open	door				打开	门			
	turn on	TV	open	TV	*	打开	电视			
	unfold	letter	open	letter	*	打开	信件			
L1<L2	wear	coat	keep	coat	*	穿	外套			
	wear	jeans	keep	jeans	*	穿	牛仔裤			
	wear	shoes	keep	shoes	*	穿	鞋			
	wear	glasses	keep	glasses	*	戴	眼镜	穿	眼镜	*
	wear	earrings	keep	earrings	*	戴	耳环	穿	耳环	*
	wear	necktie	keep	necktie	*	系	领带	穿	领带	*
	break	glasses	bring	glasses	*	打破	玻璃	违背	玻璃	*
	break	window	bring	window	*	打碎	窗户	违反	窗户	*
	break	record	bring	record	*	打破	纪录	打碎	纪录	*
	break	tradition	bring	tradition	*	打破	传统	弄坏	传统	*
	break	camera	bring	camera	*	弄坏	照相机	违反	照相机	*
	break	rule	bring	rule	*	违反	规则	打碎	规则	*
	break	promise	bring	promise	*	违背	承诺	弄坏	承诺	*
	break	heart	bring	heart	*					

Appendix F The Word Lists of Lexical Decision Task

L1-L2

	Translation		Unrelated		Related	Translation	(Nonwords)	
喝	drink	玩	drink	吃	drink	喝	drimk	*
打开	open	制作	open	关闭	open	打开	opan	*
借	borrow	发	borrow	还给	borrow	借	borraw	*
借	lend	奔	lend	归还	lend	借	lond	*
介绍	introduce	作用	introduce	说明	introduce	介绍	intraduce	*
介绍	recommend	喜欢	recommend	解释	recommend	介绍	recammind	*
看	see	抬	see	盯	see	看	sae	*
看	look	画	look	望	look	看	laok	*
看	watch	爱	watch	眼	watch	看	wetch	*
笑	laugh	摸	laugh	泪	laugh	笑	leugh	*
笑	smile	拉	smile	哭	smile	笑	smila	*
听	hear	查	hear	说	hear	听	haer	*
听	listen	考	listen	唱	listen	听	listan	*
帽子	cap	工厂	cap	围巾	cap	帽子	cep	*
帽子	hat	成绩	hat	脑袋	hat	帽子	het	*
桌子	desk	娃娃	desk	椅子	desk	桌子	dosk	*
桌子	table	银行	table	凳子	table	桌子	toble	*
高	high	美	high	低	high	高	hugh	*
高	tall	丑	tall	矮	tall	高	tull	*
大	big	新	big	小	big	大	bip	*
大	large	旧	large	面积	large	大	lurge	*
小	small	贵	small	大	small	小	smull	*
小	little	贱	little	多	little	小	littla	*
穿	wear	拿	wear	披	wear	穿	waar	*
戴	wear	证明	break	破灭	break	戴	weur	*
打破	break					打破	breok	*
弄坏	break					弄坏	broak	*
违反	break					违反	bruak	*

L2-L1

Translation		Unrelated		Related		Translation	(Nonwords)	
drink	喝	drink	玩	drink	吃	drink	噙	*
open	打开	open	制作	open	关闭	open	反开	*
borrow	借	borrow	发	borrow	还给	borrow	融坏	*
lend	借	lend	奔	lend	归还	lend	位发	*
introduce	介绍	introduce	作用	introduce	说明	introduce	违绍	*
recommend	介绍	recommend	喜欢	recommend	解释	recommend	打绍	*
see	看	see	抬	see	盯	see	坚	*
look	看	look	画	look	望	look	蝦	*
watch	看	watch	爱	watch	眼	watch	栜	*
laugh	笑	laugh	摸	laugh	泪	laugh	笔	*
smile	笑	smile	拉	smile	哭	smile	粂	*
hear	听	hear	查	hear	说	hear	峠	*
listen	听	listen	考	listen	唱	listen	晱	*
cap	帽子	cap	工厂	cap	围巾	cap	貌小	*
hat	帽子	hat	成绩	hat	脑袋	hat	冒士	*
desk	桌子	desk	娃娃	desk	椅子	desk	桌波	*
table	桌子	table	银行	table	凳子	table	卓瑞	*
high	高	high	美	high	低	high	糊	*
tall	高	tall	丑	tall	矮	tall	込	*
big	大	big	新	big	小	big	糘	*
large	大	large	旧	large	面积	large	凪	*
small	小	small	贵	small	大	small	辻	*
little	小	little	贱	little	多	little	雫	*
wear	穿	wear	拿	wear	披	wear	畑	*
wear	戴	break	证明	break	破灭	break	方打	*
break	打破					make	制习	*
break	弄坏					drive	落游	*
break	违反					play	驾生	*

Appendix G The response patterns and the RTs of “drink” and “open” for each participant in Experiment 3

G1 The response patterns and the RTs of “drink” for each participant in Experiment 3

	drink soup *	eat soup	喝 汤	吃 汤 *	drink water	eat water *	喝 水	吃 水 *	take medicine	drink medicine *	eat medicine *	喝 药	吃 药
1	E	C	E	C	C	C	C	C	C	NR	E	C	NR
1	2773	2899	1454	1487	1223	1970	993	2688	1327	0	321	1972	0
2	E	C	C	C	C	C	C	C	C	C	E	E	C
2	1186	1008	850	760	1226	1209	999	974	1299	1610	1199	756	1208
3	C	C	C	C	C	C	C	E	C	NR	E	C	C
3	2584	1855	1145	1145	1746	1254	926	1431	2025	0	2183	1322	1456
4	E	C	C	C	C	E	C	C	C	E	E	C	C
4	1346	1579	751	1233	848	845	738	1559	963	1103	1254	825	2848
5	E	E	C	C	C	E	C	C	C	E	E	C	C
5	1901	1408	954	890	956	1744	1834	1111	1070	2936	2345	1579	730
6	E	E	C	E	C	C	C	C	C	C	E	C	C
6	1092	713	849	548	65	2561	1556	736	1148	1689	1272	1144	968
7	E	E	C	C	C	E	C	E	C	E	E	C	C
7	1754	2597	539	1231	963	1467	1007	946	1701	2396	1684	782	548
8	E	C	C	C	C	E	C	C	E	E	E	C	C
8	1248	1392	906	928	1401	1511	701	869	1417	1493	1027	704	1114
9	E	E	C	C	C	C	C	C	C	C	E	C	C
9	1553	1688	1554	1009	1174	1365	623	880	1031	1616	1373	1454	2087
10	E	E	C	C	C	C	C	C	C	C	C	C	C
10	1314	1787	15	865	1409	1759	702	1052	1273	1531	1720	10	745
11	E	E	C	C	C	C	C	C	E	E	E	C	C
11	1505	1895	1151	1637	1140	1201	709	869	1617	1462	1701	965	1156

12	NR	C	C	C	C	C	C	C	C	C	E	C	C
12	0	1596	998	1469	1021	2038	1746	1108	2029	1229	1975	824	778
13	E	C	C	C	C	C	C	C	E	E	E	C	C
13	1526	1270	2090	885	1131	1349	2201	1235	2497	2161	1261	1388	1608
14	E	E	C	C	C	C	C	C	C	C	E	E	C
14	915	1048	635	650	952	847	535	988	1048	1055	694	726	569
15	E	E	C	C	C	C	C	C	C	C	C	C	E
15	1749	1317	2354	796	1515	1489	1325	864	1637	2309	1801	1248	671
16	E	C	C	C	C	C	NR	C	C	E	E	E	C
16	1741	1473	1073	940	1546	1703	0	2053	1615	1593	858	1084	1041
17	E	E	C	C	C	C	C	C	C	C	E	C	C
17	1383	1365	1004	1932	1256	1295	953	964	2657	1410	1617	1290	1180
18	C	NR	C	C	C	C	C	C	C	C	C	C	C
18	1049	0	857	857	1610	1682	1315	782	1799	1694	1551	853	714
19	E	E	C	E	C	E	C	C	C	E	E	C	C
19	978	1235	651	643	1027	1052	1044	1156	1909	1092	1260	1507	1082
20	E	E	C	C	C	C	C	C	C	E	E	C	E
20	1090	1544	743	657	1009	1311	527	928	1074	1801	1160	595	849

G2 The response patterns and the RTs of “open” for each participant in Experiment 3

	open + door	打开 + 门	open + TV *	turn on + TV	打开 + 电视	open + letter *	unfold + letter	打开 + 信件
1	C	C	C	C	C	E	C	C
1	1091	1261	288	1521	893	1805	2965	1393
2	C	C	E	C	C	E	C	C
2	969	704	832	1049	2031	982	1726	727
3	C	C	E	C	C	E	E	C
3	1395	1500	1406	1771	973	1885	2283	942
4	C	C	E	C	C	E	C	C
4	1384	1241	1294	1190	681	1132	1508	937
5	C	C	E	C	C	E	C	C
5	715	1316	880	354	1734	1300	1531	1041
6	C	C	E	C	C	C	C	C
6	1101	579	741	950	906	1312	1356	1856
7	E	C	E	C	C	E	C	C
7	82	675	850	1707	730	1644	1532	674
8	C	C	E	C	C	E	E	C
8	990	709	737	1358	649	1035	1474	776
9	C	C	E	C	C	E	C	C
9	932	1598	1005	1165	1506	1060	1499	997
10	C	C	E	C	C	E	C	C
10	906	714	966	1057	968	1333	2359	765
11	C	C	E	C	C	E	C	C
11	1012	825	1093	1248	722	1425	2192	774
12	C	C	E	C	C	E	C	C
12	1629	1378	1005	1211	845	1509	1274	1243
13	C	C	E	C	E	E	C	C

13	1683	622	654	1062	1603	1138	1462	852
14	C	C	E	C	C	E	C	C
14	1009	795	790	801	726	1055	1185	729
15	C	C	E	C	C	E	C	C
15	1424	1410	769	1635	1554	1381	2730	1410
16	C	C	C	C	C	E	C	C
16	1521	940	2018	1321	856	1709	2769	1652
17	C	C	E	C	C	NR	E	C
17	1360	1382	1049	1223	1146	0	1742	1958
18	C	C	E	C	C	C	C	NR
18	1341	2264	794	1359	1092	920	1700	0
19	C	C	E	C	C	E	C	C
19	1019	819	797	1611	1715	1066	2115	1378
20	C	C	E	E	C	E	E	C
20	1123	623	1343	2051	802	759	779	834