Effect of Word Exposure Frequency on Chinese Advanced EFL Learners’ Incidental Vocabulary Acquisition

Yeqiu Zhu
Department of English, School of Foreign Studies, Nanjing University, China

Yuxin Huang
Department of English, School of Foreign Studies, Nanjing University, China

Abstract—The present study explores the effect of word exposure frequency on Chinese advanced EFL learners’ incidental acquisition of three aspects of word knowledge (i.e., word form, word class and word meaning). The participants were 20 Chinese English postgraduates who read two chapters of an original English novel and took four vocabulary tests. The target words were 20 pseudo-words created to replace the words that naturally occurred from one to twenty times in the text. The results show that word exposure frequency has a significant effect on IVA through reading, exerting the strongest effect on word form recognition and the weakest on word meaning recall. The study also finds that seven is the threshold value for significant word gain growth and that local word frequency also influences learners’ IVA.

Index Terms—intentional vocabulary acquisition, word exposure frequency, extensive reading

I. INTRODUCTION

Incidental vocabulary acquisition (IVA) refers to the learning of vocabulary without intent to do so, as the learner’s primary objective in extensive reading is to comprehend meaning rather than learn new words (Paribakht & Wesche, 1999). The term “incidental” signifies the involuntary nature of picking up new words or word knowledge while making sense of the text. Many researchers (e.g., Day, Omura & Hiramatsu, 1991; Nation, 2001; Richardson, Morgan & Fleener, 2008) believe that incidental learning is an effective way of learning vocabulary and that L2 vocabulary is mainly acquired through incidental language learning activities, primarily through extensive reading. Nagy et al. (1987) argued that “incidental learning of words during reading may be the easiest and single most powerful means of promoting large-scale vocabulary growth” (p. 249). Krashen (1989) even claims that students can learn all the vocabulary they need from context through extensive reading.

The effect of word exposure frequency on IVA has been examined in previous theoretical and empirical studies. Researchers, however, have not yet reached a consensus on the exposure frequency for successful IVA through reading to occur. It is also noted that the majority of previous studies test word meaning recognition only and falsely equated gains in word meaning with word gains in general. Though college students primarily enlarge their vocabulary through extensive reading (Rott, 1999; Gai, 2003), many admit vocabulary acquisition is still one of their biggest problems in English learning (Meara, 1990). This study, therefore, intends to explore the effect that word exposure frequency has on Chinese advanced learners’ IVA and the threshold value for the significant effect to occur.

II. REVIEW OF PREVIOUS STUDIES

Vocabulary learning is a gradual process in which learners make word gains through repeated encounters with new words. Nation (2004) believes that one exposure to an unknown word in a text can hardly lead to vocabulary acquisition. Learners must be repeatedly exposed to a novel word in different natural contexts to achieve successful acquisition. Studies have shown that both L1 learners (Nagy et al., 1987; Shu et al., 1995) and L2 learners (Hulstijn, 1992; Horst et al., 1998; Rott, 1999) can incidentally gain knowledge of word form and meaning through reading. Jenkins, Stein and Wysocki (1984) found that word knowledge gains grew as the number of word encounters in context increased. Tekmen and Daloglu (2006) examined the relationship between learners’ IVA and word frequency in a text. They found that word frequency was a relatively strong predictor of vocabulary acquisition, accounting for 29% of the variance in acquisition.

Studies (Grabe & Stoller, 1997; Horst et al., 1998; Waring & Takaki, 2003) have also examined the effect of the vocabulary exposure frequency on IVA. The outcomes, however, varied considerably. Jenkins, Stein, and Wysocki (1984) found that about 25% of the learners could learn a word after ten exposures. Horst, Cobb and Meara (1998) explored IVA in relation to word frequency by having 34 Oman intermediate English learners read a simplified version of Mayor of Casterbridge over ten days. The students took two multiple-choice tests to see whether they could choose...
the correct meanings of the target words. The results of the study showed that words with over eight occurrences in the text were more likely to be acquired than those that occurred fewer times. Rott (1999) also examined the effect of word exposure frequency on IVA through reading. Sixty-seven German learners in the study read six paragraphs in which the target words occurred two, four, and six times, respectively. It was found that words with six exposure times led to significantly greater lexical gains. Pellicer-Sánchez and Schmitt (2010) also found that a novel word could be learned after 6 to 10 occurrences. Waring and Takaki (2003) discovered that the learners needed to encounter a word at least eight times before they could recognize its form and that they had great difficulty identifying the meaning even when the same word occurred over 18 times.

The previous empirical studies on the effect of exposure frequency on IVA lend sufficient evidence to the argument that extensive reading contributes to students’ vocabulary learning and have laid solid foundations for further studies in the field of IVA. These studies, however, are not free from limitations. The majority of previous studies tested word meaning recognition only and fail to examine the effect on word gains of different aspects separately. Lexical knowledge involves different aspects of word knowledge and should, therefore, be measured accordingly. Moreover, most studies employed multiple-choice tests and failed to measure partial word gains. Tests of this kind also allowed participants to generate answers based on their guesses instead of truly on their actual lexical knowledge. In addition, the word exposure frequencies in most previous studies were not in the order of natural word occurrences, which made it unlikely to detect the actual threshold value of word exposures.

### III. SIGNIFICANCE OF THE PRESENT STUDY

The current study aims to investigate the effect of word exposure frequency on advanced EFL learners’ incidental acquisition of word knowledge in terms of word form, word class and meaning. To address the unresolved issues, we made a stringent control of possible intervening variables. The target words in the present study were the ones that naturally occurred from one to twenty times in the reading material and were all replaced by pseudo-words to prevent the possible effect of prior knowledge of the target words. This study examined different aspects (i.e., word form, part of speech and meaning) of word knowledge and measured word meaning gains in terms of two dimensions, word meaning recall and word meaning recognition. Participants’ possible random guessing was prevented by one-on-one questioning. In this way, we can have a more accurate and in-depth understanding of the effects of word exposure frequency on different aspects of word knowledge.

### IV. METHODOLOGY

#### A. Participants

A total of 20 English postgraduate students at a Chinese key university voluntarily participated in the present study. They all have been learning English for more than ten years of English and majoring in English for more than five years. The participants included 6 males and 14 females. All of them passed with a score over 70 points (out of 100) on the Test for English Majors-Band 8 (TEM-8), the highest level of an English proficiency test specifically designed for English majors.

#### B. Reading Material

The reading material in this study was the first two chapters of the English novel *Of Mice and Men*, which none of the participants had read. The two chapters contained 12,164 words in total, with 43 words new to the participants. The percentage of new words was 0.35%, which satisfied the requirement of knowing at least 98% of the vocabulary in a text to infer the meaning of unknown words from context (Nation & Coady, 1988).

#### C. Target Words

Altogether 20 words were chosen as the target words in this study, including nine nouns, six verbs and five adjectives. These words naturally occurred in the material from one to twenty times. They were replaced by pseudo-words to eliminate the possible intervening effect of the participants’ prior knowledge of the words.

The target words were chosen according to the following criteria: 1) Words should naturally appear from one to twenty times in the text. The plural forms of a noun and the inflected forms of a regular verb (e.g., walked, walking) were counted as occurrences of the base nouns and verbs (e.g., walk); 2) Only content words were chosen, as they usually have clear referents and are more likely to be noticed and processed; 3) Words should be of different parts of speech, and nouns, verbs and adjectives should be chosen since they are the most common parts of speech in the natural text; 4) the meanings of the target words should be able to be inferred through context; 5) the meanings and parts of speech of the target words should stay the same throughout the text; 6) the pseudo-words should follow English spelling conventions and not be easily confused with real English words.

#### D. Instruments

1. Tests
This study included four tests\textsuperscript{1} that were designed to measure the effect of word exposure frequency on the acquisition of the different aspects of word knowledge, i.e., word form recognition, word class recall, word meaning recall and word meaning recognition.

Four immediate tests were carefully sequenced to minimize the effect of an earlier test on the later ones. The first test, the word form recognition multiple-choice test, was administered to see if the participants could recognize the forms of the target words. The participants chose the correct word form from the four provided items, three of which were distractors that resembled the target words in spelling. The second test, the word class recall test, was designed to measure the participants’ knowledge of parts of speech. In this test, the target words were each presented with four blanks, with no syntactic clue available. The participants would write answers based on their understanding of the word class of each target word. The third test, the word meaning recall test, examined how well the participants could infer word meanings from the context. In this test, we made twenty cards, each with one target word printed on it. The participants should explain to the researcher the meanings of the target words on the cards, and the researcher would complete the answer sheet based on the answers given by the participants. In the fourth test, the word meaning recognition test, the target words were embedded in twenty sentences. Below each of the sentences were four options of word meaning for the participants to choose from.

2. Interview
After they completed the four tests, the participants took a semi-structured interview, which helped determine partial lexical gains and the reasons for their answers to the test items.

E. Data Collection and Analysis
The experiment was carried out in a one-on-one manner with one participant at a time. The participants were informed to read the material in an hour, a time limit determined by a pilot study on two postgraduate students of English proficiency similar to that of the participants. The participants were ignorant of the research purpose and believed comprehension of the material was the focus of the task. Dictionaries were not allowed throughout the whole experiment.

Immediately after the participants finished reading, we took back the material and asked questions about the readers’ comprehension of the text. We then distributed to the participants the test papers in the following order – the word form multiple-choice test, the word class recall test, the word meaning recall test, and the word meaning recognition test. Finally, the participants took an interview, which was tape-recorded for later analysis. The whole procedure for each participant lasted about two hours.

SPSS 16.0 was used to analyze the data statistically. The descriptive statistics were obtained, and the One-way ANOVA was used to measure the effect of exposure frequency on word gains in four vocabulary tests and examine whether there existed significant differences among different frequency groups.

V. RESULTS

A. General Word Gains
The results show that incidental vocabulary acquisition can occur through reading an original English novel. The present study revealed word gains of 45.8% of the target words. The words with frequencies of 7, 10, 11, 14 and above demonstrated word gains of over 50%, which was relatively high given the incidental nature of word learning in this study. The word gains, however, were not consistent across different tests or different word knowledge aspects. The largest gain was achieved on word form recognition (60.5%), with the participants incidentally acquiring this aspect for over half of the tested target words. The second-best performance was in the word class recall test (55.5%). Word meaning recognition and word meaning recall showed lower accuracy rates, though the percentage in the former (41%) was considerably higher than that in the latter (26.25%). At the frequency of 7 (or F7), over 50% of the participants got the target words correct in the word class recall test, the word meaning recall test, and the word meaning recognition test. There existed a significant difference between word gains in the four word tests in general (p=.000) as well as significant differences between individual tests (p<.05).

B. Effects of Exposure Frequency on Word Acquisition
The present study showed that the word gains by the students substantially grew when the words occurred seven, ten and sixteen times. To identify the effect of exposure frequency on word acquisition, we divided the 20 exposure frequencies into four groups: 1-6 occurrences (Group 1), 7-10 occurrences (Group 2), 11-15 occurrences (Group 3) and 16-20 occurrences (Group 4).

\textsuperscript{1} The tests are available upon request.
Threefold. First, the aspects of target word knowledge vary across studies. The present study investigated three aspects, namely attention from text reading to word processing. Words with higher exposure frequency are processed more efficiently. There is a need to learn a new word to comprehend a text or fulfill a specific reading purpose, the learners shift their attention to word form and class. For word meaning recall and recognition tests, student performances could still exhibit visible improvement after 15 exposures. Thus, significant differences were found between Group 2 and Group 3, between Group 2 and Group 4, or between Group 3 and Group 4. This result suggested that generally speaking, words exposed over seven times were significantly better acquired and that the effect of exposure frequency was relatively weak on general word gains after six occurrences. In other words, seven exposures can be the threshold value for a new word to be substantially acquired through extensive reading.

Table 1 shows significant differences among the four frequency groups in terms of the frequency effects on target word gains, as well as between Group 1 and each of the rest groups. No significant differences were found between Group 2 and Group 3, between Group 2 and Group 3, or between Group 3 and Group 4. This result suggested that generally speaking, words exposed over seven times were significantly better acquired and that the effect of exposure frequency was relatively weak on general word gains after six occurrences. In other words, seven exposures can be the threshold value for a new word to be substantially acquired through extensive reading.

The differences among the four frequency groups were statistically significant (p<.05) in all the four vocabulary tests, which indicated the positive effect of exposure frequency on word gains in all the three aspects of word knowledge, as shown in Table 2. The frequency effect was most evident in the part-of-speech test, where the scores among four frequency groups were all significantly different (p<.05) except those between Group 2 and Group 3.

In all the four tests, words occurring over six times were significantly better acquired than those less than seven times in a text, except for between Group 1 and Group 3 in the word form recognition test. No significant differences were found between Group 2 and Group 3, which means words exposed between 7 and 10 times were not significantly better acquired than those that occurred between 11 and 15 times in a text. Significant differences, however, were found between Group 3 and Group 4 in all the four tests, which suggested that word acquisition restarted to grow at 16 repetitions after the downturn in word gains after 11 exposures. Moreover, significant differences were found between Group 2 and Group 4 in word form recognition and word-class recall tests, but not in the word meaning tests. This suggested that for word meaning acquisition, it was less likely to rely on exposure frequency to enhance performance and that for word form and class, student performances could still exhibit visible improvement after 15 exposures. Thus, seven and sixteen occurrences might both be a statistically significant indication for considerable growth in word knowledge regarding the incidental acquisition of word form and class. For word meaning recall and recognition tests, only seven can be considered the threshold value for substantial growth in understanding word meaning.

VI. Discussion

A. Frequency Effect on IVA

The present study finds that exposure frequency exerts a significant effect on IVA and that seven is the threshold value for IVA through extensive reading to occur. Kartal & Sarigul (2017) argue that input frequency significantly influences second language acquisition. This is particularly true with IVA through reading. During extensive reading, the reader’s attention is focused primarily on text comprehension, not on the form or meaning of particular words. When readers encounter a novel word, they consciously or subconsciously evaluate the salience of the word in the immediate context. A novel word remains unknown unless the reader perceives it as significant for text comprehension. Word exposure frequency to a certain extent influences readers’ judgement of word salience in the text. The greater number of word occurrence suggests the importance of words in text comprehension (Horst et al., 1998). Words with higher frequency are more likely to be perceived, especially when they are indispensable for text comprehension. Once there is a need to learn a new word to comprehend a text or fulfill a specific reading purpose, the learners shift their attention from text reading to word processing. Words with higher exposure frequency are processed more efficiently (Brysbaert et al., 2018). Consequently, IVA through reading is more likely to occur.

This study, in general, yielded a better exposure frequency effect on IVA than previous ones. The reasons are threefold. First, the aspects of target word knowledge vary across studies. The present study investigated three aspects,
i.e., word form, word class, and word meaning. Many previous studies (e.g., Horst et al. 1998) examine word meaning only, the aspect of word knowledge that is acknowledged as the most difficult to acquire accidentally. Second, the participants of this study were all English postgraduates and demonstrated higher language competence. Their higher English proficiency being developed through years of intensive English learning contributes to the positive test results of this study; even in the test of word meaning acquisition. Third, the words in the text were well controlled. Expect for the target words, over 98% of the words were familiar to the participants. This design presumably posed little difficulty comprehending the text and allowed the participants to notice and process the target words sufficiently. This design effect was confirmed by the interviews with the participants in which 90% of the participants reported their noticing of the target words.

B. Different Aspects of Word Knowledge and IVA

In terms of different aspects of word knowledge, the word exposure frequency effect is more evident in word form acquisition than in word meaning acquisition. This finding supports the argument that different aspects of word knowledge are learned at different paces (Schmitt 2000). Word knowledge, as Schmitt (2000) pointed out, consists of different aspects of lexical knowledge. Learning a word generally progresses from knowing one simpler aspect to acquiring more comprehensive knowledge. A learner’s knowledge of a word, for example, can develop from grasping orthographic and morphological features to gaining syntactic and semantic understandings. The ultimate achievement is to establish a knowledge network, in which different aspects of word knowledge connect with one another, and the newly acquired word knowledge connects with words that have already stored in one’s vocabulary reservoir. The incremental nature of word learning necessitates vocabulary research to capture the subtle word knowledge growth and to improve learners’ lexical understanding.

Word meaning comprehension and memorization are more demanding and complicated a process, and English learners generally find it most difficult to retrieve word meaning compared to the retrieval of word form and word class. The spellings and parts of speech of a word can be learned and later triggered by recognizing the orthographic form and syntactic information. However, the sight of the words and the syntactic knowledge cannot provide sufficient informative connections for meaning recall. According to the Instance-Based Theory (Bolger et al., 2008), the acquisition of word meaning is more incremental, and the core meaning of the word can only be gradually abstracted after multiple contextual encounters. Therefore, more word encounters, or higher word exposure frequencies, are needed to allow for the accumulative frequency effect on word meaning decontextualization and eventual acquisition.

C. Local Exposure Frequency and IVA

Though the participants’ accuracy rates in the tests generally increased along with the increase in word exposure frequency, there exist apparent fluctuations in the learners’ word gains. This phenomenon suggests that general exposure frequency is not the only factor that influences IVA. Local word frequency is found to exert a certain effect on IVA. Word exposure frequency, or general word frequency, generally refers to the repeated occurrence of a word in the whole text. By contrast, local word frequency refers to the number of occurrences of a word within a specific area. There are cases where words with lower general frequency but higher local frequency are better acquired than words with higher general frequency but lower local frequency. The word “badam,” for example, had a relatively higher local word frequency of 11 times and generated an acquisition percentage of 57.5%, as opposed to 50% for “bugga,” the word appearing 15 times with an uneven distribution throughout the whole text. This suggests that the local word frequency in a text can also be a factor to influence word gains.

VII. CONCLUSIONS AND IMPLICATIONS

This study investigated the effect of word exposure frequency on advanced English learners’ IVA through reading an original English novel. It is found that advanced English learners can incidentally acquire multiple aspects of word knowledge through reading authentic English materials. Of all the aspects of word knowledge examined in the current study, word form is best acquired, and word meaning recall the least. Seven is the threshold value for significant word gain growth, and sixteen occurrences witness another new high in word gains in word form and class acquisition. The results also show that word gains are not in linear proportion to exposure frequency. Local word frequency and also influences IVA.

The findings of the present study have some practical implications for vocabulary learning. It is observed that few advanced English learners intentionally memorize new words, as many believe it is time-consuming and inefficient. The present study shows that English learners of higher proficiency can learn words by reading long authentic English materials and that word reoccurrence increases the chances of a word being noticed and processed. Extensive reading, therefore, can be an effective way to enlarge learners’ vocabulary size. To achieve the optimal vocabulary acquisition effect, English learners need to read materials long enough to ensure that learners can encounter the same new words at least seven times, the time that generates the significant frequency effect to occur. It should also be noted that word learning involves different aspects of word knowledge, which often develop at different paces. Frequency effect is most evident on word form recognition and least on word meaning recall. Learners should learn to understand that learners can hardly fully acquire a word with contextual word learning and that only unbalanced word knowledge is gained.
through extensive reading. Therefore, acquisition of word meaning also entails conscious effort on their part to learn the correct meaning from the context and retain the knowledge.

REFERENCES


Yeqiu Zhu received her PhD in English Language and Literature at Nanjing University, China. She is currently an associate professor at the School of Foreign Studies at Nanjing University in Jiangsu, China. Her research interests include second language acquisition, second language pedagogy, cognitive linguistics, and corpus linguistics.

Yuxin Huang holds an MA in English Language and Literature. Her research interests include second language acquisition and applied linguistics.