Revisiting Word Exposure Frequency and Incidental Vocabulary Acquisition*

Yeqiuzhu
English Department, School of Foreign Studies, Nanjing University, Nanjing, Jiangsu, China

Abstract—Incidental vocabulary acquisition (IVA) has long been an important, yet controversial topic in second language acquisition. Word exposure frequency is one of the key factors that influence IVA. Various studies have been conducted to explore exposure frequency in relation to IVA, and have not yet reached an agreement on the effect and the optimal exposure frequency. The exploration of the existing studies reveals the reasons for the conflicting results, which include insufficient reading amount, simplification of test content and lack of consideration of learners’ lexical competence. To clear up the blurred picture and better understand the nature of word exposure frequency effect, future researches can be carried out in terms of research areas and methods. In terms of research areas, studies can be designed to test different aspects of word knowledge, distinguish the effect of global word frequency and local word frequency, and investigate such mediating effects as perceptual salience, and skill levels. Methodologically, the test format needs to be diversified to avoid pure multiple choice tests. In addition, more post-tests can be conducted, and qualitative and longitudinal studies should be employed as complementary to the popular quantitative ones.

Index Terms—word exposure frequency, incidental vocabulary acquisition, reading

I. INTRODUCTION

Incidental vocabulary acquisition (IVA) is believed to occur mostly through reading, especially extensive reading, as extensive reading provides readers with precious opportunities to encounter and process unfamiliar words in natural contexts. Word exposure frequency is one of the key factors that influence IVA. Various studies have been conducted to explore exposure frequency in relation to IVA, and have not reached a consensus on the effect and the optimal exposure frequency. IVA through extensive reading is generally regarded as a good way to enlarge language learners’ vocabulary size, but it is also argued that word knowledge gains from IVA as a result of multiple encounter with words can be inconsequential enough to overlook, as evidenced by some existing studies. IVA is an important, yet controversial topic worth exploring and cannot be easily discarded without thorough investigation. To understand the effect of word exposure on IVA through extensive reading and learn about its optimal frequency, one should acknowledge the complexity of the learning process of IVA, and give serious consideration to such sophisticated aspects involving in IVA as word knowledge, word gains, the amount of input (i.e., the amount of reading materials learners are exposed to) and time over which learners are reading. This article is intended to explore the concept and the features of IVA, evaluate the existing studies to explain the reasons for the conflicting results and discuss the research areas and methods worth exploring in future studies.

II. UNDERSTANDING IVA

The concept of incidental vocabulary acquisition, first proposed by Nagy, Herman and Anderson in 1985, refers to “learning of vocabulary as the by-product of activity not explicitly geared to vocabulary learning (Hulstijn, 2001, p.264).” It is a type of learning with no intent to learn vocabulary when the learner’s primary objective is to comprehend or communicate, rather than to learn the unknown word itself. As a concept put forward as opposed to intentional vocabulary acquisition, “incidental” in incidental vocabulary acquisition signifies the involuntary nature of “picking up” new words or word knowledge while trying to make sense of the text and/or to complete other required tasks.

Related to the concept of IVA is intentional vocabulary learning, which refers to the learning of vocabulary by deliberately committing lexical information to memory. In intentional vocabulary learning, learners focus their attention on words in various exercises and activities. Such exercises and activities include learning words with annotations or vocabulary cards, playing word games, etc. Intentional learning, when learners have an evident intention of learning and retaining lexical information, often activates conscious processing of words, which is often realized through the use of techniques like rehearsal and memorizing techniques (Nation, 2004).

The essential distinction between incidental and intentional vocabulary learning in terms of psychological processing lies in the learner’s initial “intention” or “consciousness.” In contrast to intentional word learning as a lexicon-oriented learning with a direct and explicit focus on words and expressions, IVA is a meaning-oriented learning outcome, mainly

* Supported by a project (11YJC740165) of the Ministry of Education, China, entitled “Noun Acquisition in different languages: Interlanguage development from the perspective of conceptualization.”
resulted from extensive reading (Stanovich, 1986; Krashen, 1989; Nagy et al., 1987). Memorization of a new word, in this sense, comes as a natural result of this process. IVA is closely related to the textual context whose richness of information can lead to successful inference of word meaning and acquisition of different aspects of word knowledge (e.g., word form, word meaning, word class, etc.). One obvious advantage of IVA through extensive reading is that a large amount of reading significantly increases learners’ encounter with unknown words, and provide more word learning opportunities for inferencing.

III. IVA AND WORD EXPOSURE FREQUENCY: A CONTROVERSIAL ISSUE

Exposure frequency influences word gains in two ways: one is by increasing the possibility of a word to be noticed and processed by the reader, and the other is by strengthening the association between lexical stimuli (target words) and mental responses (cognitive processing). A word of high frequency in the text, therefore, is likely to receive more attention from readers, as a frequent occurrence of a word demonstrates itself the importance in text comprehension. As a result, words exposed repeatedly in a text tend to evoke more cognitive processing and are easier to be stored, activated and retrieved later on. Word exposure frequency, among all the factors that influence IVA through reading, is the best studied. Though IVA and word exposure frequency are arguably correlated, researchers have not reached a consensus on 1) how effectively word exposure frequency can enhance IVA, and 2) how many times of exposure are needed for successful IVA through reading to occur.

Various studies have been conducted to explore the effect of word exposure frequency on IVA, and all agree that one exposure of an unknown word in a text can hardly lead to vocabulary acquisition. The acquisition percentage of the words occurring only once in a text is around 0.05% (Nagy et al., 1987). To achieve successful word acquisition, learners must be exposed to an unknown word repeatedly in a various natural contexts.

L2 vocabulary is believed to be acquired mostly through incidental language-learning activities, especially through extensive reading. Nagy et al. (1987) believes 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 will learn all the words they need from context by reading extensively. One important reason for this claim is that IVA is closely related to the textual context whose richness of information may lead to successful inference of word meaning and acquisition of other aspects of word knowledge. While reading materials, learners learn the meanings and connotations of new words, and, with considerable accumulative amount of reading, how to use the words in certain contexts, even though their original purpose for reading is not to learn words.

Incidental vocabulary acquisition through reading is believed to be a gradual process in which gains are made in small increments with repeated encounters. Waring and Takaki (2003) proved the possibility of IVA through reading, and showed higher acquisition and retention rates among words of frequent occurrence. The authors suggested that there was a 50% chance of correctly identifying word form in the post-test three months later if learners met the word eight times or more. Tekmen and Daloglu (2006) examined the predicting power of word frequency in a text over word learning. Three groups of Turkish learners of English participated in the reading of an authentic text (The Golden Fleece) with about 2,400 words and four lexical tests administered over a two-week period. The results indicated that frequency was a fairly strong predictor of IVA, with 29% of the variance accounted for by exposure frequency.

Some other researchers, however, are far less optimistic, as their studies produce less positive results. Horst, Cobb and Meara (1998) asked 34 intermediate L2 learners to read a simplified Mayor of Casterbridge (109 pages) over a ten-day period, and found that the general acquisition rate was 20% for all subjects. Similarly, Jenkins et al. (1984) found in their study that only about 25% of the learners acquiring a word after 10 times of exposure. Wu and Xu (2006) conducted a research specifically on the effect of word frequency on vocabulary acquisition and gained similar results. Fifty five Chinese non-English majors in college participated in the reading of an English story containing 3,332 words, after which a multiple-choice test and a blank-filling test on target-word knowledge were administered. The results showed that the general acquisition rate was about 20% in the two tests, and that readers might incidentally acquire a word through reading when it occurred 14-17 times in the text. Zhang and Qi (2009) combined qualitative and quantitative research methods and investigated whether seven weeks of extensive reading improved four participants’ two aspects of word knowledge: spelling and meaning. The study yielded detailed, yet discouraging results: extensive reading could facilitate IVA with an average acquisition rate of 21.6% for spelling and 15.5% for meaning.

The other controversial issue concerning IVA is the threshold value of word exposure frequency if the frequency indeed exerts positive effect on vocabulary acquisition. A study conducted by Saragi et al. (1978) found out that the subjects learned most of the words that were presented to them six times or more, but words that were exposed fewer than six times were learned only by half of the subjects. Horst, Cobb and Meara (1998) explored IVA in relation to word frequency and learner vocabulary size. The results of their study showed that words which occurred over 8 times in text were more likely to be acquired than words that were exposed less. Wu and Xu found that readers incidentally acquired a word through reading when it occurred 14-17 times in the text.

© 2015 ACADEMY PUBLICATION
Instead of viewing word gains as a whole, some researches investigate the effect of word exposure frequency on different aspects of word knowledge. The results indicate that the influence of exposure frequency varies across knowledge types. Orthographic knowledge benefits greatly from three times of exposure but not from an additional four and for semantic knowledge, most gains occur between three and seven times (Chen & Truscott, 2010). Waring and Takaki (2003), however, discovered that the learners needed to encounter a word at least eight times before they could recognize its form, and in an unprompted meaning test, learners were confronted with great difficulty while identifying the meaning even though the word was exposed over 18 times.

IV. MEDIATING FACTORS: UNDERSTANDING THE CONTROVERSY

The existing researches have provided abundant evidence as to whether and how extensive reading helps improve students’ vocabulary learning. Concerning the effect, more and more researches support the positive effect of word exposure frequency over IVA, though some remain suspicious of the inconsequential exposure effect. The more thorny and complex issue is exactly how many times learners need to encounter a new word before they can acquire it. The factors that affect the outcomes of IVA through reading are multitudinous, including vocabulary size, and text and word characteristics (Swanborn & De Glopper, 1999). With multiple factors, together with word exposure frequency, co-functioning in IVA, the lack of control of certain variables and/or insufficient consideration of certain mediating factors in the previous studies explains the reasons why the researches fail to reach a consensus on the effect of the word exposure and times of exposure.

First of all, the generally poor results of IVA concerning word exposure effect may be due to subjects’ insufficient reading amount. Vocabulary cannot be fully acquired without sufficient amount of input and learners’ conscious processing of morphological and semantic information of a word. Not all the new words learners encounter can get into their lexicon. Only those words they consider important in or relevant to text comprehension are processed with the help of linguistic, interlinguistic and extralinguistic information. During the process of extensive reading, the reader’s attention is focused primarily on text comprehension, not on the form or meaning of particular words. When encountering an unfamiliar word, the reader will consciously or subconsciously evaluate how central the word is in the immediate context. The unfamiliar word would be perceived as unknown and then processed usually when it causes comprehension breakdown. Once there is a need to understand an unknown word in order to comprehend a text or fulfill certain reading purposes, the reader will direct his/her attention to the word. The stronger the need is, the more likely a word is to be noticed and the deeper it is to be processed. It is observed, however, that most of the reading materials used in previous studies contained only a few thousand words and the so-called extensive reading in experiments lasted for several weeks at most. The unknown words might not have enough exposure frequency or input data for readers to notice and process, thus possibly impairing the effect of repeated exposure.

Second, the intricacies of the target words are not well controlled or explicitly indicated. Words are of different nature in learning difficulty and characteristics, and cannot be treated as equal in studies. Laufer (1997) points out that some words are easier to learn and others are more difficult. But many existing researches neglect this difference, testing words as a whole and failing to indicate the difficulty level of target words. It is highly possible that words were tested in different studies are on different difficulty levels, thus yielding contradicting results. Also overlooked are intralexical characteristics of a word like word class, context richness, grammatical function, etc. All this co-functions to affect learners’ word acquisition, but in most, if not all, studies, researchers considered the number of occurrences of target words alone.

In addition, test items are often limited to word form/meaning recognition, which falsely equated gains in word form/meaning with word acquisition. Lexical knowledge is far from an all-or-nothing phenomenon, but involves various aspects of knowledge. Nation (2001) proposed a word knowledge classification, listing the aspects of both receptive and productive word knowledge, including form (sound, spelling, and word structure), meaning (associations, referents, and the concept expressed), and use (the patterns a word appears in, its collocations, and constraints on its use). More aspects of word knowledge need to be measured in separate tests. What makes tests even questionable is the fact that in most studies, only multiple-choice tests are employed, which is by no means sufficient enough to measure subtle improvements on word understanding and partial word gains.

Finally, learners’ lexical competence often keeps unknown, which might result in the inaccurate measurement of word exposure effect on IVA. To make adequate word gains through reading, the reader must know 90 to 95% of the words in a given text. Learners’ prior word knowledge relates not only to learners’ minimum vocabulary size, but to the comprehensibility of a given text. Correct inference of word meanings is conditional upon accurate recognition of surrounding words in the context. If it is not carefully controlled, learners’ language proficiency can be an influential intervening factor in the measurement of exposure frequency effect, which would make it difficult to justifiably claim the presence or absence of word exposure frequency effect. Studies should either control or measure the effect of language proficiency, and ignoring it is the last option to choose.

V. EXPLORING IVA IN FUTURE STUDIES: POSSIBLE AREAS AND METHODS

The abundant research results, though sometimes conflicting, have laid solid ground and provided valuable
implications for further studies in the field of IVA through reading. Future studies are needed to probe deeper into the issue and they can be carried out in terms of two aspects: research areas and methods.

A. In Terms of Areas

Future studies can measure more aspects of word knowledge, in addition to the well-studied aspects: word form and meaning. Among the aspects of word knowledge proposed by Nation (2001), apart from sound, which is hardly possible to pick up accurately through reading, the acquisition of all the other aspects can be achieved and measured in one way or another. The majority of the existing studies, if not all, probe into the receptive word knowledge and leave the productive one untouched. Though it usually takes more time to learn the effect of incidental acquisition of productive word knowledge through reading, this is not entirely impossible and is well worth investigating. In addition, more studies can be conducted to find out the effect of word exposure frequency on word class acquisition. Content words and function words carry different weight for the necessary understanding of a sentence, and learners acquire them at different paces. The same is true with different kinds of content words. Nouns, for instance, are believed to be better acquired than verbs. But this claim awaits more empirical evidence to verify. Besides, little is known about word gains of adjectives and adverbs. In this sense, more empirical studies are needed to test the findings of the existing studies, and to explore and compare word gains of other different word classes. Due to different cognitive involvement loads, learning gains may not be consistent across aspects of word knowledge. The larger gains would be most probably on word form recognition and word class, which is less cognitively demanding than the other aspects, as our processing of information moves from a sensory level of analysis, through pattern recognition, syntactic realization, to semantic enrichment. The memorization of the formal aspects, e.g., word spelling and part of speech, stays on the initial level of registering the visual images of words and the environment where they occur, which are easier to be triggered by the sight of the words. By contrast, word meaning acquisition could be the hardest of all, because reading words alone might not provide sufficient associations or information needed for meaning acquisition.

Word knowledge includes not only receptive one, but productive one as well. Nearly all the existing studies focus on the former, and neglect the latter. One generally needs less information about a word to recognize it than it does to generate it. Measurements for productive competence, such as translation tests and sentence-making tasks, can be used to gain a whole picture of how well learners retain and use words, and to compare between receptive and productive word knowledge in terms of the effect that exposure frequency has on word gains and retention.

In addition, future researches can identify three types of word frequency (i.e., general word frequency, global word frequency, and local word frequency) and investigate and compare the effects of these three. General word frequency refers to the frequency of a word that occurs in English in general, global word frequency is the total amount of a word that appears in the whole text, and local word frequency is the number of occurrences of a word within a clustered area. Words of high frequency in general do not necessarily guarantee word gains. It is learners’ encounter frequency with words that counts. Words of high frequency in English are likely to be poorly acquired if their local word frequency (three times on average) is low. In the same line, words with low general frequency but high local frequency could be better acquired than words with high general frequency but low local frequency. Based on these findings, studies can go further and dig deeper to find out how frequently target words should be globally and locally arranged in a text in order to maximize the learning effect.

While frequent encounter with a word in different contexts certainly enhances the chance of picking up correct word forms, making close judgments on word classes, and inferring word meaning, exposure frequency alone might not ensure higher acquisition rates of unknown words. The chance of a word to be processed and remembered largely depends on the possibility of it being noticed. Word exposure frequency is only one of the significant determinants of noticeability identified by Schmidt (1990). The other determinants related to IVA are perceptual salience and skill levels. These two, together with word exposure frequency, all exert influence over how effectively learners can incidentally acquire words new to them. Therefore, apart from the direct correlation between word exposure frequency and IVA, equally worth investigating are the mediating effects of the two above-mentioned factors and how they co-function with word exposure frequency.

Word salience is crucial for readers to notice and then process a word new to them. Noticing is more of personal choice, which depends on the learner’s assessment of target words’ importance in the context. Words that are selectively attended to and carefully processed have a bigger chance to stay in the long term memory. Context is paramount in the process of integrating word knowledge into learners’ lexical system. The context where an unknown word appears must provide adequate information for the reader to guess its meaning in the absence of word lists, dictionaries or any other external assistance. If the word is only peripherally significant in understanding the text, the reader is most likely to ignore the word. In Paribakht and Wesche’s research (1999), learners skipped approximately half the words they identify as unknown in text reading. The more important a word is considered and the deeper it is processed by the reader, the more successful the IVA of the word is likely to occur. The knowledge about the underlying structure of a complex stimulus environment has to be triggered in the exact same or at least a highly familiar context. Learners’ inference is key to IVA, which asks for both learners’ ability and context richness to ensure IVA. Learners make reasonable guesses on the meaning of an unknown word with all available linguistic cues from the context in combination with the learner’s inter- and extra-linguistic knowledge. It is, therefore, essential to understand in what context new words can more salient to learners and learners are more likely to focus on word processing.
In any empirical study, there exist different independent variables. The effect of one factor on the dependent variable can be difficult to measure when intervening variables function simultaneously. There is no exception for the study of IVA. The factors that influence IVA, other than word exposure frequency, include learners’ prior knowledge of the words, lexical competence, word inference difficulty, etc. Researchers, therefore, need to well control all the other factors to minimize their mediating effect on the target factor. Besides, learner’s deliberate or accidental encounter of the target words could also be an intervening factor. To preclude post-experiment encounters, pseudo-words can be designed in studies. To achieve best resemblance to natural words, pseudo-words should be carefully designed and go through a pilot study to avoid any morphological confusions to the subjects.

B. In Terms of Methods

Methodologically speaking, the study of IVA should include different types of test. When there are multiple choice tests only, as in many existing studies, test takers can make a wild guess and have an at least 25% chance of success. Such random guessing can be prevented by one-on-one questioning and answering or learners’ self-assessment of their word learning. Researchers need to gain more specific and detailed data concerning how well learners can truly learn the words new to them. In addition, two measurements at different sensitivities (recall and recognition) can be employed to assess word meaning gains. In this way, a more accurate assessment of the degrees of IVA and partial word gains can be achieved and the effects of frequency on different aspects of word knowledge can be investigated respectively. The accuracy rate of word recognition is argued to be considerably higher than that of word recall, but more studies are needed to verify the claim.

Besides, both quantitative and qualitative analyses can be adopted. The statistics generated from quantitative analyses help find out the general patterns of the effect of word exposure frequency, and qualitative data yield more detailed and specific information. Previous studies, mainly based on collected quantitative data, have yielded some insightful results, but it is observed that some hold conflicting views, and some give presumptive reasons accounting for phenomena or patterns emerging from the quantitative analyses. There is no easy way to truly understand in what way and to what extent learners process and acquire word knowledge, unless researchers can find some hard evidence. Qualitative analyses can be used as a complement to quantitative ones. Interviews with subjects, for example, can provide more thorough and subtle analyses of the IVA progress and minor gains that may not be captured by vocabulary tests.

Longitudinal studies (i.e., arranging more exposure time for learners to read) can also be conducted to test the effect of word frequency. The existing studies usually test the effect after several hours’ reading, seven weeks at most. Recalling any word learning experiences through reading, we find that the amount of time which extensive reading invests is crucial for word gains and for the development of many other language-related skills as well. No one can considerably expand his/her vocabulary size overnight. It usually takes months, even years to notice visible progress. The lack of sufficient reading time might be the reason why some studies argue the effect of word exposure frequency on IVA is so trivial and insignificant that it can be neglected. A longitudinal study of one year or more will probably reveal some insightful findings that have had been veiled.

Last but not least, the delayed posttest can be designed in a more refined way to test the retention of more aspects of word knowledge over different periods of time. The existing studies usually examine the retention of word meaning only. Since different types of word knowledge require different encoding processes and may be retained for different periods of time, future studies can measure the retention of various aspects of word knowledge and probe into the effect of exposure frequency on IVA with respect to each aspect.

VI. Conclusion

IVA has long been a controversial, yet important topic in second language acquisition. Researchers have not come to the agreements as to the significance of word exposure frequency effects on IVA and the threshold value of exposure frequency that predicts substantial word gains. Various careful studies have been conducted to probe into these issues and presented important, insightful findings. Yet, a comprehensive review of these studies reveals such problems as the short time span of the experiments, the simplification of test content and the casual choice of subjects. All this may well explain the inconsistence in the findings.

To clear up the blurred picture and better understand the nature of word exposure frequency effect, future researches can be carried out in two main aspects: research areas and methods. In terms of research areas, studies can be designed to test different aspects of word knowledge, distinguish the effect of global word frequency and local word frequency, and investigate such mediating effects as perceptual salience, and skill levels. Methodologically, the test format needs to be diversified to avoid pure multiple choice tests, more post-tests can be conducted, and qualitative and longitudinal studies should be employed as complementary to the popular quantitative ones.

Vocabulary acquisition is believed to be one of the fundamental aspects in developing language proficiency. Though this view is widely shared, many language learners find enlarging vocabulary size rather demanding and cannot be easily accomplished. Few learners, especially advanced learners, intentionally memorize new word lists, as they view this as an unproductive, time-consuming task. Practical experience and previous studies have proved the feasibility of learning new words through reading. Teachers and learners both have learned extensive reading can be an effective, though probably time-consuming way to considerably enlarge their vocabulary size and acquire word knowledge. The
existing studies provide solid evidence and insightful findings concerning the effect of word exposure frequency on IVA. Answers have been found to questions like to what extent and in what way IVA helps language learners pick up words new to them in reading contexts and how long this effect can last. Future studies that are more sophisticated and refined in research areas and design can describe and explain in depth how word exposure frequency influences IVA, co-functions with other factors, and provide more practical and specific suggestions as to how to learn words effectively through reading.

REFERENCES


Yeqi Zhu received her Ph.D. in English Language and Literature, in Nanjing University, China. She is currently an assistant professor in School of Foreign Studies at Nanjing University in Jiangsu, China. Her research interests include second language acquisition, second language pedagogy and corpus linguistics.