Considerations when researching EFL vocabulary acquisition through Extensive Reading

Thomas BOUTORWICK*

Introduction

This article will first detail a hypothetical study which will set the context to explain some of the considerations that should be remembered when conducting research into English as a foreign language (EFL) learners’ vocabulary acquisition through extensive reading. Before the research explanation, a brief outline of relevant terms will be detailed.

Relevant Research

Incidental vocabulary acquisition

Incidental vocabulary acquisition (IVA) can be defined as “... a by-product of language usage, without the intended purpose of learning a particular linguistic feature” (Schmitt, 2010). This byproduct of language usage can occur when, for example, when reading a novel simply for pleasure. The key point of IVA is that the focus of attention is on the meaning of the text. This contrasts with a course textbook, for example, in that when reading a textbook, the focus is also on the learning of a new linguistic feature. These features can be anything from orthography to collocation knowledge of a word.

Extensive reading

Extensive reading (ER) involves reading text, whether a novel or a short piece of literature, with a focus on the meaning of the text. (Nation, 2001, p. 149) It should be clear that ER and IVA go hand in hand as the prerequisites for ER allow

* Instructor of English as a Foreign Language (IEFL), Department of Science and Technology, Kwansei Gakuin University
for IVA to occur. One question that appears often in the literature is that of quantifying ‘extensive’. In other words, what exactly constitutes extensive? At the time of writing this, there is no number that accurately quantifies ‘extensive’.

**Productive vocabulary knowledge**

One major phenomenon in second language vocabulary acquisition is the distinction between receptive knowledge and productive knowledge of a word. Receptive knowledge of a word refers to the ability to recognize a word’s form and meaning link when, for example, a multiple choice test is given. In a multiple choice format test, a word is prompted and the test taker must choose the correct answer, or definition, of that word. In this case, all necessary information is provided; the word’s form and its meaning, and it is up to the test taker to recognize that link. In contrast, productive knowledge refers to one’s ability to produce a word, without the aforementioned clues. This means two main things: First, it means that productive knowledge of a word is more difficult to master. It also means that productive knowledge is harder to predict; there are many ways to say the same thing in a language, and as such the act of not using a word doesn’t necessarily mean lack of knowledge.

**Theoretical study**

What follows is an outline of a research proposal. All of the theory is based off of previous research, i.e. Webb (2002) and Waring & Takaki (2003). The following research questions were devised for this study, which in turn will set the stage for some considerations to take into account for this area of research.

**Research Questions**

1. Does Extensive reading of ten graded readers lead to an increase in
   a. Productive cued-recall of a word’s form?
   b. Productive cued-recall of a word’s form-meaning?
2. Is there a correlation between a word’s frequency of occurrence and productive knowledge acquisition of that word?
3. Is productive knowledge of nouns acquired more than productive knowledge of verbs?

The hypotheses for each research question were:

1. Extensive reading of ten graded readers does lead to an increase in:
   a. Productive cued-recall of a word’s form
   b. Productive cued-recall of a word’s form-meaning
2. There is a positive correlation between a word’s frequency of occurrence and productive acquisition of that word.
3. Productive knowledge of nouns is acquired more than productive knowledge of verbs.

All of these hypotheses were formed based off of previous research of which the main areas relevant to this study have been explained in the previous section.

Method

This research was conducted at a private university in Japan. Ninety-six (96) first-year and second-year students, aged 18–20 years old, participated in the study. Over the course of ten weeks, the students read ten graded readers (GRs) at the pace of one graded reader a week. After reading each graded reader, they composed a short summary of the book as well as a response to the book. The response that the students wrote included information such as the degree of difficulty of the book, the enjoyment level of the book, etc. At the end of the 10 weeks of reading, the students were given two tests that assessed their vocabulary knowledge of some of the words that were found in each book.

Graded Readers

There were ten graded readers used in the study. All of the books were part of the Penguin Readers series. Of the ten books, 5 were level 0 (starter) and 5 were level 1 (beginner). The books are written in British English, and the starter books contain 200 headwords, while the beginner books have 300 headwords. A headword in this case refers to a word family, or in other terms all of the inflections and derivations of a word. According to this definition, the words “acted”, “actor”, and “acting” would all be considered part of the same word family of “act”.

The average score for the first 1000 words was 8.1 out of 10, equating to roughly 810 words. This estimate of knowledge may not be high enough to provide the 95% coverage rate that the literature has recommended (Nation & Webb, 2011). To that extent, level 0 graded readers were incorporated in the study as they provide fewer headwords, equating to a higher lexical coverage rate. Relatedly, 10 level 0 GRs were not used because with approximately 200 headwords in a level 0 book, and previous research showing that at least 95% of the words should be known in order for incidental learning to occur, the amount of headwords that would need to be substituted would be 10 (200 * 95%). This is of course assuming a 100% coverage rate. This is a very small amount of words to test, most likely leading to results that would not be able to be generalized to the population, and so level 1
GRs were also used to increase the headwords and thus increase the amount of words that could be tested while maintaining a 95% or higher coverage rate.

**Target Words**

Twenty-nine (29) target words were used in this study. Of the 29 words, 20 were nouns and 9 were verbs. These words were replaced with substitute words that were taken from Clements-Stephens et al. (2010). Each of the substitute words were real words with the exception of one letter being replaced. Also, all of the words were monosyllabic. All of the words substituted were from the first 1000-word band of the British National Corpus. The target words were chosen solely from the first BNC word band to maximize the likelihood that students would already be familiar with the concept that the word refers to.

The literature has suggested that there should be roughly a six to four ratio of nouns to verbs, as this maintains authenticity; it mirrors how the language actually occurs. With 29 target words that means roughly 17 nouns and 12 verbs. In this study however, words were selected based not only on part of speech, but also range, frequency of occurrence and level according to the British National Corpus (i.e. the first 1000 frequency band). It was thus impossible to keep this ratio. This resulted in 20 nouns and 9 verbs.

**Reading**

At the beginning of each class, for 10 weeks, each student was given the same book. They were told to read and try to understand the meaning of the story. They were also told that they could not use any means to assist in their learning, i.e., a dictionary, the Internet, etc., because the purpose of the activity was to enjoy the story and afterwards write a short (3–5 sentence) summary and additionally a response about the book.

After they finished reading, they gave the book back to the teacher and then opened their computer where they found the form that they were to fill out. The form provided a way to assess whether students had actually read the book or not. Second, it was a possible avenue for them to use the target words in a free-recall fashion, showing that productive word knowledge acquisition took place. It was emphasized that the summary of the story was not as important as the response, for two reasons. First, it was assumed that prolonged recall of the story would force explicit vocabulary recall. Second, the form was part of the extensive writing section of the course which emphasizes fluency over accuracy.

Before the students read each book, the form was explained so that they would be focused on understanding the meaning of the story. They were not told during the ten weeks that they would be tested on vocabulary at the end of the final book.
Considerations

Now that the groundwork has been laid out, some of the considerations can now be discussed.

Testing considerations

When testing productive knowledge, it is very likely that a floor effect will occur in a test unless measures are taken to prevent it. This is because research seems to suggest that IVA, while leading to increases in knowledge, has been found to have small effect sizes. As such, piloting the test before administering it can lead to results that are reliable and valid.

Another similar idea is that of previous knowledge. It is impossible to know exactly what words a person knows, especially when considering the depth of knowledge that exists with every word. As such, the best one can do is to try to conduct diagnostic tests to assess the participants in a study.

Methodological considerations

In the study above, one of the main limitations is that there was no control group. It is thus difficult to say, given the results, whether ER lead to acquisition of a word’s form, or if it was just student knowledge that lead to them decoding the pronunciation. This is where a control group would be ideal so that a comparison could be made between the experiment group, the group that did the reading, and the control group that was not exposed to the reading. In this manner, it would possible to rule out the easily recognizable factor of orthographic knowledge of English.

Regarding ER, there are two considerations as they relate to the above study. The first is the idea of quantifying ‘extensive’. Is reading one book considered extensive? If not, then is 10 considered a good number to qualify as extensive? It could be that there is no concrete number, but a relative number depending on the group that is doing the reading. If this were the case, then comparisons could be made between groups to determine what was causing one group to need less exposure for more gains. In other words, this would allow for the optimization of ER so that maximum gains could be made.

The second consideration with this study is the criteria for ER. The above definition suggests ‘simply for pleasure’ as one of the main requirements for ER to occur. In the above study however, the short reports that were written for each GR were counted as part of the course grade. As such, it is possible that the students were reading more in-depth than if they would for pleasure reading. The extent of the differences in reading between these two conditions has yet to be tested, and it
may be assumed that there would be minimal differences, however good research involves mimicking reality to the best of one’s ability.

**Definition of a word**

One of the most fundamental aspects of doing researching into the effects of ER on vocabulary acquisition is the selection of target words. Deciding which words to test should be based on the purpose of the experiment. This may be an obvious statement, however it can be broken down into different areas that, when taken as a whole, reveal the depth of thought that must go into deciding words to test. First, one must decide what is considered a word. Nation (2001) proposes that a lemma be used as the unit of measurement for a word when conducting research into productive knowledge. A lemma in this case refers to a word and all of its inflections. In other words, a word that, when inflected keeps the same part of speech. This seems like a valid method for choosing a definition of a word because the word form that a student may need to use is partly decided on the part of speech that is needed.

**Conclusion**

Research into ER is beneficial because it has a direct relation to learning and teaching a foreign language. Understanding how reading interacts with vocabulary acquisition can help to answer questions regarding the ideal frequency of occurrence figures for IVA to take place, and the kinds of words (i.e., noun, verb, etc) that are most effectively learned incidentally. It should be clear that vocabulary acquisition is not an simple dichotomy. One simply does not know or not know a word. There are many factors that affect learning a word, and numerous aspects of word knowledge. Through principled, empirically based research however, it will be possible to make more sense of this multi-dimensional phenomenon that is the building block of language learning.

**References**


Palgrave Macmillan.