RELATIONSHIP BETWEEN READING COMPREHENSION AND VOCABULARY SIZE AMONG ENGLISH MAJORS AT THE UNIVERSITY OF APPLIED SCIENCES OF KAUNAS

Vaida Miseviciute[1], Vita Speckauskiene[2]

[1]University of Applied Sciences of Kaunas, Lithuania

[2]Lithuanian University of Health Sciences, Kaunas, Lithuania

Abstract

Reading is one of the most important language skills that helps build students' vocabulary. Even though many other skills are involved in increasing reading comprehension, vocabulary size may be a significant aid in improving understanding of an academic text. The purpose of the study is to examine the relationship between students' reading comprehension skills against their vocabulary size. Additionally, the reading comprehension scores at the entry level are compared to the reading comprehension scores after reading instruction at midterm point. A total of 11 first year students seeking BA in business English and enrolled in the intensive English course "English for social communication" at the University of Applied Sciences in Kaunas participated in the study. Statistical data analysis using chi-square test was used to analyze independence between two qualitative (nominal) variables—the productive and receptive reading comprehension test scores at the entry level and midterm point against the vocabulary size test. The findings indicate the students with larger vocabulary have better reading comprehension, yet teaching students to read effectively and building reading skills allow students with smaller vocabulary to have equally good comprehension of the reading both receptively and productively. The findings reveal the importance of developing students' reading skills and vocabulary and should motivate students to seek to understand and manipulate available language rather than just memorize new vocabulary words.

Keywords: *vocabulary size, reading proficiency, receptive and productive vocabulary*

INTRODUCTION

Reading is one of the most essential ways in which student vocabulary can be built. Because reading provides incidental vocabulary learning, reading plays a crucial role in the acquisition of language. Unfortunately, social media, where reading is limited to deciphering texting codes, emoji language, or snippets of essential information, is slowly altering reading habits among the young. To form an opinion, develop critical thinking, understand an issue at its core, reading long academic texts is essential in the complex world today, yet during the intensive English classroom, the impression starts forming that the students are scanning long texts rather than reading. When longer texts are assigned, there is that deep sigh of torture. Scanning the text just to grasp the idea may be enough for some purposes, but it is often not enough in an academic setting. Notably, reading is essential in an ESL classroom where reading aids language acquisition. This hypothesis is being tested to understand the relationship between the reading and vocabulary size among students. Even though research points to the complexity of reading and vocabulary relationships, the purpose of the study is to analyze and better understand how the vocabulary size influences receptive and productive type reading questionst the entry level and midterm point. It is believed that vocabulary knowledge is a vital language learning component in reading comprehension. Of course, paralinguistic abilities like previous knowledge and the ability to make connections between what is known and new, or the knowledge of actively adapting reading skills such as ability to understand inferences or main ideas, as well as direct linguistic knowledge of grammar and sentence structures, ability to understand referents, detect explanations of the new words contribute immensely towards understanding a reading. Nevertheless, there is a clear connection between the vocabulary size and reading. According to several research studies, 95% of the words must be familiar to the reader to understand an academic text (Hirsh & Nation, 1992; Laufer, 2003).

VOCABULARY SIZE MEASUREMENTS

There are many ways to measure the vocabulary. The very first attempts to measure student vocabulary were carried out by the teachers asking students to mark unknown words in the dictionary pages. For example, if a student was given a 10,000 word dictionary and did not know 50 words out of 100 words on random dictionary pages, the conclusion was made, that the student know 50% of the words and had a 5,000 vocabulary. A more accurate attempt at calculating the vocabulary size was proposed by Paul Nation. The original version of 14,000 vocabulary size test version contained 140 multiple-choice items, with 10 items from each 1000 word family level. In order to estimate the total receptive vocabulary, the total score was multiplied by 100. This means the test started by testing words at the 1,000 frequency level, then the 2,000, then the 3,000 and so on up to 14,000 frequency level (Nation & Beglar, 2007). Other, more accurate vocabulary count tests were designed, yet according to Paul Nation the 140 item test works just as efficiently as it has a good balance of items from various frequency levels (Nation, 2012)

Regarding the efficiency of the test, Beglar's (2010) concluded that 140 item vocabulary size test was accurately measuring receptive vocabulary knowledge, and it distinguished between learners of different proficiency levels. One of the most important messages that teachers need to understand is that the test is measuring only written receptive vocabulary knowledge. Paul Nation clarifies that the test does not measure speaking or listening vocabulary, nor does it measure the vocabulary needed to writing or reading texts. This naturally begs the question why bother testing the receptive vocabulary size, and how it could affect instruction. In fact, knowing the vocabulary size of language learners is important for setting goals in classroom programs (Nation & Webb, 2011).

There are many different ways in which the vocabulary size measurement could be used. Firstly, it may provide some understanding about the learner's abilities to perform certain tasks. Several observations were made regarding reading comprehension and fluency having examined research findings by Nguyen, & Nation, 2011; Nation, 2012; Wren (2000). It is possible to see the connections between the vocabulary size and the abilities in table 1. Table one has been constructed by the author to help visualize the abilities based on the native speaker and a second language learner abilities.

Age	**Native speaker**	**Abilities**	**ESL learners**
4 year old	1,500	Basic conversation survival skills	
10 year old	5,000	General conversation Simplified/adapted readings	5,000-8,000* Most speakers Enough vocabulary to comprehend readings
High school grads	5,000 -10,000*	Fluent intelligent conversation Good reading comprehension	8,000-15,000* ESL speakers who have lived in English speaking countries
BA	15,000-20,000*	Academic work/discussions Full reading comprehension	
MA	20,000-30,000*	Complex academic work Full comprehension of complex academic readings	

* regular readers who read 2-4 hours per day

Table 1. Vocabulary size measurement statistics (prepared by author)

The table clearly shows that the majority of the second language learners have the vocabulary size of 5,000 -8,000 word families and this corresponds to a 10 year old native speaker who can have a fluent conversation and read simplified or adapted texts with full comprehension. Of course, second language learners who read a lot or who have a chance to study or live in an English speaking country may easily develop a vocabulary up to 15,000 word families, and such level of vocabulary allows for successful academic work. Interestingly and relevant here is the discrepancy in the reading comprehension findings. While judging native speaker reading comprehension, full comprehension is achieved when the vocabulary size reaches 10,000-15,000 word families, yet when talking about ESL learners the numbers are reduced to 5,000-8,000 word families. This is very understandable since according to many researchers most of the ESL learners can only reach this vocabulary size.

Secondly, the vocabulary size measurements may be useful for placing ESL students into levels or assigning textbooks. In fact, at the earlier stages of the development the Common European Framework of Reference for Languages (CEFR) vocabulary size measurements were included, yet in 2003 such scales were removed and replaced by the skills based and knowledge based criteria (Council of Europe, represented in the article by Meara and Milton (2003) in table 2.

CEF level	*Cambridge exam*	*XLex score (max. 5000)*
A1	Starters, Movers and Flyers	<1500
A2	Kernel English Test (KET)	1500–2500
B1	Preliminary English Test (PET)	2750–3250
B2	First Certificate in English (FCE)	3250–3750
C1	Cambridge Advanced English (CAE)	3750–4500
C2	Cambridge Proficiency in English (CPE)	4500–5000

Table 2. CERF level, Cambridge exam and vocabulary size (from Meara and Milton, 2003, p. 5)

The results in this table show that a vocabulary of 4500-5000 word families is enough to reach C2 level of English, and therefore, perform academic tasks; however, many studies that were carried out doubted that this amount of vocabulary is enough to perform fairly complex academic tasks that are required at a C2 level. For example, Milton and Hopkins (2006) criticized the 4,500 – 5000 word families rage for the highest level (C2) on the Common European Framework of Reference (CEFR) reading though as the learners would not be able to achieve the skills listed in the descriptors. Much higher vocabulary size numbers were suggested in Hu and Nation (2000) and Nation (2006) who estimate 8,000 – 9,000 word families for learners to be able to read texts like novels and newspapers without any assistance (dictionary, instructor, etc). Similarly, Schmitt (2008) concluded that while it is possible to understand texts with a smaller vocabulary, around 8,000 – 9,000 word families are required to read authentic texts at a high comprehension level. A later study of 661 participants found a linear relationship between the vocabulary knowledge and reading comprehension (Schmitt, Xiang Ying &, Grabe, 2011).

Therefore, this knowledge may aid in choosing the right textbook, reading materials, or an academic vocabulary list that is the most relevant for the classroom. Finally, a vocabulary size measurement may be used as a motivator for students to strive to know more or to build confidence that they have already reached a lot.

VOCABULARY ACQUISITION

There are two ways in which vocabulary can be acquired—deliberate and incidental. Deliberate learning occurs through direct vocabulary instruction. Direct attention to vocabulary is especially important at the low proficiency levels and studies show significant benefits in deliberate instruction (Scarcella & Oxford, 1992). According to Nation (2001), the, one of the most influencial researchers on the vocabulary studies, the first 2,000 most frequent words in addition to the words on the Academic Word List merit intentional instruction due to the utility of knowing common and useful words. Another strong opinion is expressed by Schmitt (2008)--a direct vocabulary approach "always leads to greater and faster gains, with a better chance of retention and of reaching productive level of mastery" (341). However, every practitioner or language learner will agree that vocabulary learning is not that simple. Lists of vocabulary words that start piling up quickly, even in the most exciting ESL classrooms, often feel decontextualized and overwhelming. Along with the strong opinions that direct vocabulary instruction works, there are just as strong opponents who claim that deliberate way of learning vocabulary is the least effective way of learning vocabulary (Hulstijn &Laufer, 2001; Notion 2006; Schmitt, 2008; Dodigovic, Jeaco & Wei, 2017).

Another type of learning is incidental learning. This type of vocabulary learning occurs while speaking/listening (learning from peers, teachers, speakers) or reading. Vocabulary is being picked up naturally and effortlessly. Nevertheless, incidental learning is more likely to occur when students have a high-proficiency level and might read for pleasure (Nation, 2001). The findings about incidental learning are just as conflicting. On one hand, reading is often found to be the most effective incidental way of learning vocabulary (Luo, 2013; Ramos &Dario, 2015). On the other hand, the studies carried out by Waring & Takaki (2003) and Pigada & Schmitt (2006) have found relatively low pick up rates for learning new words incidentally. Such confusion in the research results simply points to the fact that learning and teaching vocabulary is complex and depends on many variables.

To name a few, here are the more commonly mentioned conditions that seem to be effective when teaching/learning vocabulary: Huckin & Coady (1999) caution that "reading… does not automatically lead to the acquisition of vocabulary. Much depends on the context … and learner's attention" (p. 183). On a similar note Duke& Bennett-Armistead (2003) stress that learning occurs when teachers "immerse students in words in a variety of ways and get them personally and actively involved in constructing word meanings" (p. 182) and Kandybovich (2017) notes that students remember words better when the readings or themes are interesting. Since everyone has different interests, many authors have noted the importance of the exposure to rich and various context and themes (Nation, 2006; Dodigovic, 2015). Gardner (2008), for example suggests reading novels or texts of similar themes/ similar vocabulary to build repetition in various contexts. Regarding repetition, studies pointed to different number of exposure to new vocabulary necessary. Perez &Alvira (2017) analyzed the studies and discovered the number of encounters reported ranged from five to sixteen (Nation, 1990), six (Rott, 1999) or more (Thornbury, 2004), at least eight (Waring & Takaki, 2003), and more than ten (Pigada & Schmitt, 2006).

However, today's research focuses more on smart repetition and such trend can be observed in vocabulary acquisition since 1990's. This is due to the complexity theory that was brought to language acquisition by Diane Larsen Freeman and the concept of affordance that calls for heavy repetition in various contexts, but also deeper studies into memory starting with Piotr Wozniak's study of super memo first published in 1994. He proposed a complex computerized system based on an algorithm to enhance memory. A complex calculation helps learners recycle words just at the right time. The goal to make vocabulary learning as efficient as possible resulted in the Spaced Repetition System (SSR)--if you review information too early, you'll waste your time. If you review information too late, you will forget it and will have to re-learn it. Such conclusions were proposed by Wixted & Ebbesen's (1997) research and Ebbinghaus's forgetting curve claiming that we forget 58% of what we learn after the first 20 minutes, and only 33% of the information remain after one day. The same findings were supported by Nation (2001) who agreed that a forgotten word is a lost time investment.

In summarizing the two learning ways of vocabulary and the current suggestions on the best practices for learning, it is clear that vocabulary teaching is complex and depends on many variables. Even

though Nation (2007) in his research proposes to use the principle of the four strands that calls for a well balanced and equal amount of time spent on each component in the classroom: meaning focused input, meaning focused output, language focused learning, and fluency development, she concluded that the respondents seemed to overemphasize language-focused learning at the expense of fluency development. The analysis of the data further stressed the belief that language-focused learning was seen as more efficient means of teaching and fluency was mere fine tuning and a priority. Similar observations were made by Brown (2020) who speculated that the perceived interest in grammar may be due to the disconnect between teaching and testing. Students preferred language focused learning and specifically grammar to communicative practice and even valued those teachers who explained, practiced, taught language directly.

It is important to note that while there is a disagreement about what type of readings and what teaching methods may be the most effective in building vocabulary, none of the researchers question the value of reading. It is important to stress to the students that reading is one of the most essential ways of building vocabulary and their language proficiency even if they tend to value grammar over other language skills.

READING COMPREHENSION AND VOCABULARY

To begin with, it is important to acknowledge that vocabulary alone is not enough to comprehend the text. Firstly, it is important to consider paralinguistic abilities such as previous knowledge, background information, and the ability to make connections between previous knowledge and other readings and new information, ability to understand what is important or not in the text in reading. Additionally, reading skills such as detection of main ideas and supporting information, knowing when to scan or skim the text, how to look for clues, understand inferences, sarcasm, opposing view, etc all contribute towards reading comprehension. Moreover, the subtle differences in grammar tenses or various sentence structures that may alter the meaning of the text cannot be ignored when talking about full text comprehension. Finally, it is the reading skills regarding vocabulary itself-- the skills to guess the vocabulary in context, the derivational knowledge of Latin and Greek stems, the ability to understand referents, or a simple awareness that some more complex words are given direct explanations by the authors themselves. All of these skills may influence hugely the ability to comprehend a reading and altogether prove that vocabulary solely may not be the most important requirement for effective reading.

On the other hand, many scholars argue that without the comprehension of the majority of the vocabulary in any given text, it is merely impossible to employ all the other reading skills. One of the most important considerations in dealing with reading comprehension is the density of the unknown words a reader can tolerate. According to Hu and Nation (2000) who measured reading comprehension using multiple choice questions, the learners needed to understand at least 98% of the text to answer the questions correctly. This can be translated to a density of 1 in 50 unknown words. Similar findings were made by Schmitt, Xiang Ying and Grabe (2011) who conducted a study on 661 participants and found that there is a linear relationship between vocabulary growth and learner's text comprehension. Their study also highlights the importance of the threshold level for text comprehension. These research findings provide evidence that higher vocabulary knowledge leads to higher text comprehension. Furthermore, the research findings conclude that the 5,000-8,000 range is enough to understand fairly complex readings at a C1 level. However, these results beg the question how well the readings are understood and how the reading questions can be answered receptively (where the students only choose the correct answer versus productively (where the students have to write their answer by using paraphrasing techniques. Ultimately, the purpose of the study is to clarify how the knowledge of the students' vocabulary size can help teachers understand the students better and inform instruction (specifically, vocabulary or reading).

RESEARCH QUESTIONS

In this research study, the vocabulary size of the first year students enrolled in the BA Business English philology program at the University of Applied Sciences in Kaunas was investigated, and the relationship between reading comprehension of academic texts and vocabulary size was examined. The following questions were raised: 1) What is the vocabulary size of the first year students at the entry of the program? 2) What is the relationship between reading comprehension and vocabulary size of the first year students at the entry test? 3) What is the relationship between reading comprehension and vocabulary size of the first year students after intensive reading skills instruction?

METHODOLOGY

Participants. The study was carried out at a public University of Applied Sciences with the first year students who just started their BA philology studies in Business English. The study was carried out with 15 students enrolled in an intensive (6 hours per week) English course "English for Social Communication." Students followed integrated language skills curriculum, yet special attention was paid to vocabulary acquisition and reading comprehension during the first half of the course. The students in this study were tested two times. Within the first week of the course, the students were asked to take an entry level test and a vocabulary size test (vocabularysize.com). At the midterm point (week 8 of the course), the students had to take a midterm exam. 11 students completed all the tests. Four students did not participate in the study because they did not record their vocabulary size or did not take either one of the tests.

To complete the statistical analysis, the students were grouped into 2 groups based on their vocabulary size. The decision to split the students into two groups was made due to a small participant sample. The logic behind the split was explained in Table 1, where the first group represents the majority of the non –native speakers with the vocabulary of 5,000-8,000 word families, and the second group contains the students with a larger vocabulary. In this study, it represents the students with the vocabulary of 9,001 and over 14,000 word families. For convenience, the groups will be called the Lower Vocabulary Group (LVG) and the Higher Vocabulary Group (HVG).

The reading test. The two reading tests were used from the teacher's book designed by the Oxenden &Latham-Koenig's textbook New English File (Advanced) meaning that the readings and the tests were designed for the students seeking to reach C1 level by CERF. Since the tests integrated all language skills (reading, writing, and listening), only the reading portion of the test was used to record vocabulary size and reading comprehension scores.

The reading tests consisted of two different tasks: receptive and productive tasks. Receptive tasks did not ask students to produce any language. Receptive type questions included several multiple choice questions based on the details, facts, or opinions from the two page reading as well as a few matching questions where the students had to match the main idea to the given paragraph of that reading. Several questions were used to test students' productive skills. The students were asked to write their own responses to the questions by paraphrasing information they received in the texts. The productive questions asked students to paraphrase the information or connect the reading to the previous knowledge in order to test full comprehension of the text. During each of the reading tests, 7 receptive and 7 productive type questions were asked.

Vocabulary size test. Vocabularysize.com test was used in the study. According to the *Vocabulary Size Test* information page, the test is based on the principles of Paul Nation test and cannot measure all aspects of vocabulary correctly. The creators cautioned that all the tests are flawed in that they measure certain things only—recognition of the word, meaning of the word, or use of the word. Each question may be formatted to test one or another subtle information about the word; therefore, it is important to acknowledge the underlying assumptions, construction, and format of vocabulary tests can result in different aspects of the word knowledge which the test measures (Nation, 2001). In fact, one of the goals of VocabularySize.com is to assist researchers in piloting, validating, and distributing vocabulary tests which measure a range of different aspects of word knowledge. The test development

information can be found in the article "A Vocabulary Size Test" (Nation & Beglar, 2007). For the purpose of this study, vocabulary size refers to the word families as defined by Nation (2001: 8) which consists of "...a headword, its inflected forms, and its closely related derived form." This means that if the student can recognize the word beautiful, the student is assumed to know all the other words with the same head (beauty, beautifully, beautify…).

Statistical data analysis was performed by IBM SPSS Statistics 27. Qualitative data was presented by frequencies and relative frequencies in percentage. Chi-square test was used to analyze independence between two qualitative (nominal) variables. The relation was estimated by calculating Cramer's coefficient of association. Relation between variables considered to be statistically significant if $p<0.05$.

FINDINGS AND DISCUSSION

The findings are presented based on the three research questions that were raised: 1) What is the vocabulary size of the first year students at the entry of the program? 2) Is there a difference between receptive and productive skills on a reading test at the entry level? 3) What is the relationship between reading comprehension and vocabulary size of the first year students at the entry test? 4) Is there a difference between receptive and productive skills on a reading test after intense reading skills instruction? 5) What is the relationship between reading comprehension and vocabulary size of the first year students after intensive reading skills instruction?

Research question 1. What is the vocabulary size of the first year students at the entry of the program? Table 3 represents reported student knowledge of the vocabulary size.

Vocabulary size (word families)	Number of students
5,000-7,000	2
7,001-8,000	1
8,001-9,000	0
9,001-10,100	1
10,101-12,000	2
12,001-13,000	3
13,001-14,000	0
more than 14,000	3

Table 3. Representation of vocabulary size of the students

The results regarding the vocabulary size were surprising because based on the multiple studies, the vocabulary size of the majority of the second language learners is predicted to be between 5,000-8,000 word families. The fact that 8 out of 12 students have a vocabulary equaling to the vocabulary size of the native speakers at the BA level is astounding. Granted, these students are English majors, yet the vocabulary size is very high.

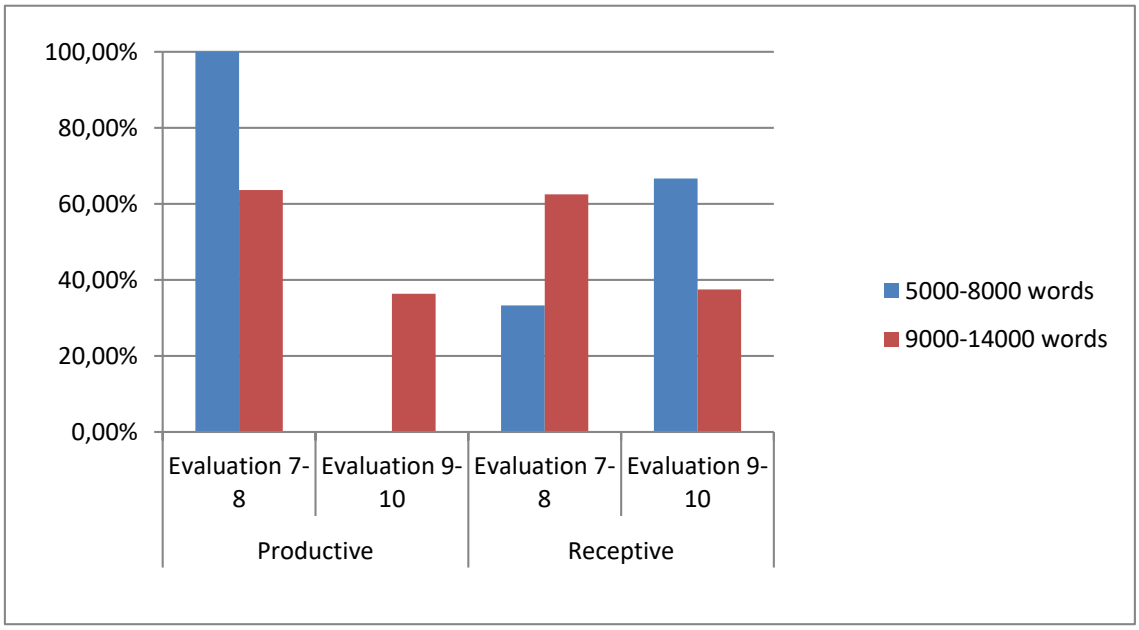
When thinking about this number several explanations may be possible. Students in today's world have a high exposure to English. Even if they do not have a chance to live in an English speaking country, they are exposed to authentic language through music (songs), TV (movies, documentaries, shows, cartoons, etc), social media, and last but not least gaming. This exposure to native speaker English at a fairly young age plays an important role in the vocabulary size. Many research studies support the fact that the early student exposure to a foreign language through instruction, television or media increases their language skills (Lambine (2008); Haznedar&Uysal, 2010; Webb, 2010;

Ghaderpanahi, 2012; Al Zoubi, 2018). Considering that the selected student population has been fascinated with languages and planned their career with English may explain their interest in acquiring English vocabulary.

On the other hand, the result brings a lot of suspicion because it is hard to believe that 50% of the participants have the vocabulary that is as large as the native speaker vocabulary after graduation. Therefore, the ESL vocabulary is most likely receptive. While the students can recognize the words, it is most likely that they cannot use the words in production. Regarding this observation, Paul Nation (2012) clarifies that the test does not measure speaking or listening vocabulary, nor does it measure the vocabulary needed to write or read authentic texts. It is important to note that the active (productive) vocabulary consists of the words/word families that students understand, recall words on demand in speaking and writing with either correct spelling or pronunciation. The passive vocabulary is recognized and understood in some contexts and may not be understood or used in other ones. Therefore, while it is a huge confidence booster for the students, it is important to clarify to students that the high vocabulary size number does not indicate successful production, and students need to continue to acquire vocabulary by applying various vocabulary learning techniques. Since vocabulary size testing provides such obscure information, this may naturally beg the question why bother testing the receptive vocabulary size. According to Nation & Webb (2011), knowing the vocabulary size of language learners is important for setting goals in classroom programs (Nation & Webb, 2011). On one hand, the teachers may make better choices what level of vocabulary to tackle (e.g. 5,000 word level, 8,000 word level, or 10, 000 word level). On the other hand, the students may understand where they are at in terms of language learning. One of the questions that the study aims to answer is the vocabulary size connection to the reading comprehension.

Research question 2. What is the relationship between reading comprehension and vocabulary size of the first year students at the entry test?

During the first week of the course, students are given a reading test consisting of two different reading comprehension questions: 7 receptive type questions and 7 productive type questions. All the students passed the test and received high grades. Two groups were separated for discussion: those who received an excellent evaluation (grade 9 or 10) and those who received good evaluation (grade 7 or 8). The difference between the receptive and productive skills on a reading test at the entry level is represented in Graph 1.



Graph 1. Evaluation received on receptive and productive reading skills at the entry level

Regarding the productive reading skills, the results show that students in the High Vocabulary Group (HVG) receive better grades than the students at the entry level test in the Lower Vocabulary Group (LVG). In the HVG, 38% of the students received a perfect grade (9 or 10) and 62% of the students received a good grade (7 or 8). Respectively, the students in the LVG did not receive perfect grades (neither 9 nor 10). The productive skills questions are harder because the students must show not only comprehension of the text, but also display paraphrasing skills that require higher vocabulary production skills. The results clearly show that the students in a HVG are more able to express their answers and score higher at a reading test regarding productive skills. No statistically significant relations were found between these evaluations.

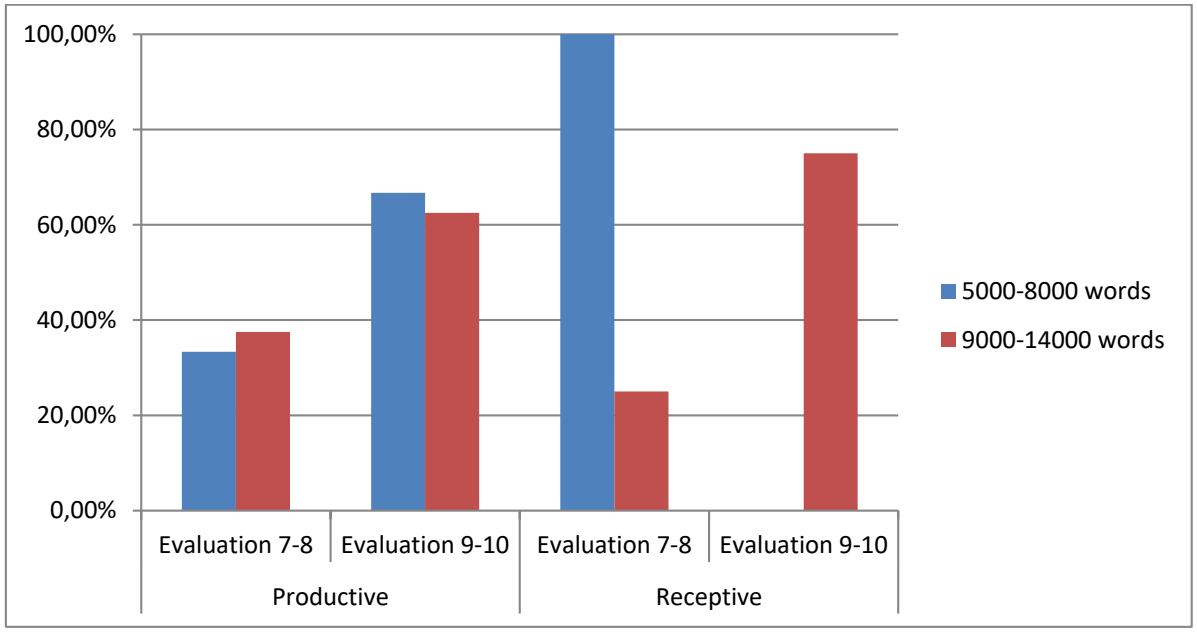
Regarding the receptive skills, the results seem highly surprising because the students in the HVG received lower scores. Only 38% of the students in HVG received a perfect score and 62% received a good grade. Respectively, while 64% of the students in the LVG received a perfect score and 36% received a good grade. This result is surprising because higher vocabulary naturally suggests better comprehension. This contradicts the results of Schmitt, Xiang Ying and Grabe (2011) where a linear progression is seen between vocabulary size and the reading comprehension. However, this result may also be due to a fairly small size of the participants in the study or a high level of vocabulary size that the students are starting with. According to the Hu and Nation (2000) the students with 5,000-8,000 vocabulary may display full comprehension, and since the skills are only receptive—the students simply have to match the answers or choose a correct answer. This result may show that 5000-8000 size may be adequate in comprehending the text and performing the tasks productively. Students have enough vocabulary at this level to answer the questions in their own words and capture the meaning of the reading. This finding corresponds to the findings by Schmitt, Xiang Ying and Grabe (2011) who claimed that this is a threshold level for text comprehension, and it is enough to tackle complex texts.

Research question 3. What is the relationship between reading comprehension and vocabulary size of the first year students after intensive reading skills instruction?

During the first half of the semester, a very special attention is given to effective reading skills. During each lecture, a skill is introduced and enforced by practice exercises. The following skills are being taught:

1) Smart reading (skills for scanning, skimming, perusing; reading speed; chunking)
2) Main idea (direct or indirect, typical location, implied main ideas, construction of main ideas (focused topic+ controlling idea). The practice starts with multiple choice questions and proceeds into forming individual main ideas.
3) Organization (typical rhetorical modes (description, narration, cause and effect, argumentation, comparison/contrast and transitions or specific words connected to each mode)
4) Cohesion (sentence level (given-new information principle), paragraph level, whole text level)
5) Summarizing (from main ideas to supporting information)
6) Vocabulary (referents, guessing unknown words, academic word list versus general word list, technical vocabulary, stems and affixes)
7) Grammar (simple, compound and complex sentences, and difficulties specifically experienced by Lithuanian students that cause comprehension issues)
8) Critical thinking and logics in reading (confirmation bias, heuristics, inferences, hidden information)

At midterm point, the student reading comprehension skills have been tested. Similarly to the entry test, the midterm exam reading section contained two different reading comprehension questions: 7 receptive and 7 productive type questions. All the students passed the test and received grades between 7 and 10. Two groups were analyzed—those who received an excellent evaluation (grade 9 or 10) and those who received good evaluation (grade 7 or 8). The difference between the receptive and productive skills on a reading test at the midterm point is represented in Graph 2.



Graph 2. Evaluation received on receptive and productive reading skills at the midterm point

Regarding productive skills, the LVG and HVG groups scored about the same. 61% of the students in HVG and 63% of the students in LVG received a perfect score. Respectively, 39% and 37% of the students received a good score. This result is interesting as it shows the importance of reading instruction. In the HVG, the number of high scoring students increased from 38% to 61%. In the LVG, the number of high scoring students increased from 0% to 63%. This shows not only that the students score higher when they are given reading instruction, but it also shows that the students in LVG can progress even faster than the students in the HVG.

Regarding receptive skills, 78% of the students in the HVG received a perfect score and 22% of the students in HVG received a good score. Students in the LVG did not receive a perfect score. All of the participants received a good score. A more predictable result is seen from the receptive skills. The LVG scored significantly lower than the HVG in the receptive portion of the test. This helps clarify the findings of Hu and Nation (2000) who used multiple choice questions to understand the relationship between reading comprehension and vocabulary size. According to them, the learners needed to understand at least 98% of the text (1 unknown word out of 50) to answer the questions correctly. Apparently, when talking about multiple choice questions that are based on the understanding of the details, facts, or the opinions of the reading text, the vocabulary size is more significant because the one unknown word may actually alter the meaning and comprehension of the text. Even through the students are instructed on how to read better, the receptive questions often test the understanding of the key vocabulary on which the detail, the fact or an opinion may depend. Of course, it is important not to get carried away with these observations not only because the grammar knowledge or reading skills may impact the results, but also the test taking skills and the guessing in multiple choice questions may skewer the results. Nevertheless, the result here is significant in that it send a very clear message that the students in the HVG have better reading comprehension that the students in the LVG. However, no statistically significant results were found.

Comparing the results from the entry level to midterm point, the results provide some interesting insights. In HVG, more students received a perfect score when analyzing productive skills. It increased from 38% to 78%. This may again be due to the instruction. Analyzing receptive skills, in the LVG, 64% of the students received a perfect score at the entry test whereas at midterm point nobody received a perfect score. There is a huge decrease from the entry level to midterm point. This result goes against expectation. There are several possible explanations for this. First, it is important to consider that the test contains only 7 receptive questions and a random text that may be harder/easier

for some students based on their previous knowledge. Secondly, it is important to note that the text and the questions are also harder as the textbooks are written to increase students' English skills. Another important point is the instruction. The more or better students understand the reading skills and the reading, the less they guess. It is important to consider that receptive questions are often answered based on a good guess rather than full comprehension of the text.

CONCLUSIONS

The aim of the study was to examine the relationship between reading comprehension skills and the vocabulary size of the students majoring in English. The following conclusions can be made:

Firstly, most of the students majoring in English have a higher vocabulary than 9, 000 word families which exceeds the vocabulary range set for the students in previous research studies. Secondly, 5, 000-8,000 vocabulary size is enough to show high reading comprehension. This result corresponds to previous findings as over 60% of students in LVG received a perfect score when receptive reading skills were tested.

Thirdly, the students in a HVG receive better scores when productive reading skills are tested. Vocabulary size matters when students are asked to produce and answer and not merely make a choice on a multiple choice question.

Fourth, the students score higher when they are given reading instruction. Furthermore, students with LVG can progress faster and demonstrate better reading comprehension when reading instruction is provided. This is an important finding as reading skills can be taught faster whereas building a higher vocabulary may take much longer.

An overall analysis of the results shows that while the vocabulary size is important and may aid reading comprehension and help students achieve better results, after the threshold of 5,000-8,000 word families is achieved, full text comprehension may be reached due paralinguistic factors.

This may be a good motivator for the students to continue to build their vocabulary as well as a great piece of information for teachers regarding the vocabulary level that needs to be addressed in the classroom. Although the nature of the research sample is very small and focuses on English majors, these conclusions may provide some general pedagogical implications. Specifically, it is essential to highlight the 5,000-8,000 vocabulary size threshold needed for full reading comprehension. There seems to be a positive correlation between reading skills instruction and the reading comprehension. If the students have lower vocabulary, yet are taught the reading skills, their reading comprehension scores increase. Building academic vocabulary eases understanding and reading comprehension. It becomes easier to critically assess and evaluate information and survive in today's world.

REFERENCES

Al Zoubi, S. M. (2018). The impact of exposure to English language on language acquisition. *Journal of Applied Linguistics and Language Research*, *5*(4), 151-162.

Beglar, D. (2010) A Rasch-based validation of the Vocabulary Size Test. Language Testing, 27(1), 101-118.

Brown, A. V. (2020). Students' and teachers' perceptions of effective foreign language teaching: A comparison of ideals. *The Modern Language Journal*, *93*(1), 46-60.

Dodigovic, M., Jeaco, S., & Wei, R. (2017). Trends in Vocabulary Research. *TESOL International Journal*, *12*(1), 1-6.

Duke, N. K., & Bennett-Armistead, V. S. (2003). "Reading & writing informational text in the primary grades."New York, NY: Scholastic.

Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. Language Learning, 61(2), 367-413

Gardner, D. (2008). Vocabulary recycling in children's authentic reading materials: A corpus-based investigation of narrow reading.

Ghaderpanahi, L (2012). Using Authentic Aural Materials to Develop Listening Comprehensionin the EFL Classroom. English Language Teaching Journal Vol. 5, No. 6

Haznedar B. & Uysal H. H. (2010). Handbook for teaching foreign languages to young learners in primary schools (pp. 1-20). Ankara: AnıPublications.

Hu, M. & Nation, P. (2000). Unknown Vocabulary Density and Reading Comprehension. Reading in a Foreign Language. 13.

Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language: A review. *Studies in second language acquisition*, *21*(2), 181-193.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language learning*, *51*(3), 539-558.

Kandybovich, S. (2017). Ten ways to learn new words as a language learner. British Council. Retrieved on June 6 at https://www.britishcouncil.org/voices-magazine/ten-ways-learn-new-words-language-learner

Lambine, C. (2008). English just isn't a foreign language anymore. Språk og språkundervisning, 2, 5-7.

Laufer, B. (2003). Vocabulary Acquisition in a Second Language: Do Learners Really Acquire Most Vocabulary by Reading? Some Empirical Evidence. Canadian Modern Language Review-revue Canadienne Des Langues Vivantes - CAN MOD LANG REV. 59. 567-587. 10.3138/cmlr.59.4.567.

Luo, J. P. (2013). Is Incidental Vocabulary Acquisition Feasible to EFL Learning?. *English Language Teaching*, *6*(10), 245-251.

Meara, P. & Milton, J. (2003). X_Lex, The Swansea Levels Test. Newbury: Express.

Milton J. & Alexiou, T. (2009). Vocabulary size and the Common European Framework of Reference in Languages. In B.Richards, H. Daller, D. Malvern, P. Meara, J. Milton and J. Treffers-Daller (eds), Vocabulary studies in first and second language acquisition. Palgrave: Macmillan, 194-211.

Milton, J. & Hopkins, N. (2006). Comparing phonological and orthographic vocabulary size: do vocabulary tests underestimate the knowledge of some learners. The Canadian Modern Language Review, 63(1),127-147.

Nation, I. S. P. (2001) *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Nation, I. S. P. (2006). How Large a Vocabulary Is Needed for Reading and Listening?. Canadian Modern Language Review-revue Canadienne Des Langues Vivantes - CAN MOD LANG REV. 63. 59-81. 10.1353/cml.2006.0049.

Nation, I. S. P. (2007) The Four Strands, Innovation in Language Learning and Teaching, 1:1, 2-13, DOI: 10.2167/illt039.0

Nation, I.S.P. (2012). A vocabulary size test. Victoria University of Wellington. Retrieved on April 25, 2021 from https://www.wgtn.ac.nz › the-vocabulary-size-test

Nation, I.S.P. & Beglar, D. (2007) A vocabulary size test. The Language Teacher,31(7),9-13.

Nation, I. S. P., & Webb, S. A. (2011). Researching and Analyzing Vocabulary. Heinle, Cengage Learning.

Nguyen, L. T. C., & Nation, I. S. P. (2011). A bilingual vocabulary size test of English for Vietnamese learners. RELC Journal, 42(1), 86-99.

Pérez, L. & Alvira, R. (2017). The Acquisition of Vocabulary Through Three Memory Strategies. Colomb. Appl. Linguist. J.,19(1), pp. 103-116.

Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a foreign language*, *18*(1), 1-28.

Ramos, R., & Dario, F. (2015). Incidental vocabulary learning in second language acquisition: A literature review. *Profile Issues in TeachersProfessional Development*, *17*(1), 157-166.

Scarcella, R. C., & Oxford, R. L. (1992). The tapestry of language learning: The individual in the communicative classroom.

Schmitt N. (2008). Review article: Instructed second language vocabulary learning. Language Teaching Research. 2008;12(3):329-363. doi:10.1177/1362168808089921

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, *95*(1), 26-43.

Ramonda, K. (2012). The four strands of vocabulary learning: Reaching a crossroads of practice and research. Vocabulary Education & Research Bulletin. 1. 5-7.

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader?.

Webb, S. (2010). A corpus driven study of the potential for vocabulary learning through watchingmovies. International Journal of Corpus Linguistics, 15(4), 497-519.

Wixted, J. & Ebbesen, E. (1997). Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions. Memory & cognition. 25. 731-9. 10.3758/BF03211316.

Wren, S. (2000) The Cognitive Foundations of Learning to Read: A Framework. Southwest Educational Development Laboratory, Austin.