

The Role of Vocabulary Breadth and Depth in Reading Comprehension: A Quantitative Study of Finnish EFL Learners

Noora Harkio

Department of English, University of Turku, Finland

Päivi Pietilä

Department of English, University of Turku, Finland

Abstract—This article reports the results of a study on the relationship between second language vocabulary breadth, vocabulary depth, and reading comprehension. A special aim was to discover the role of vocabulary depth in the prediction and explanation of L2 learners' reading comprehension. Two proficiency levels, intermediate and advanced, were compared. Vocabulary breadth was measured with the Vocabulary Size Test (Nation & Beglar, 2007), vocabulary depth with the Vocabulary Knowledge Scale (Wesche & Paribakht, 1996), and reading comprehension with a test compiled of sections from two former matriculation examination tasks. The three constructs showed strong positive correlations in both groups of subjects. However, based on the results, vocabulary breadth and depth seem to be stronger predictors of reading comprehension skills in lower levels of proficiency than on the advanced level.

Index Terms—L2 English, vocabulary breadth, vocabulary depth, reading comprehension, language learning

I. INTRODUCTION

The importance of vocabulary knowledge for reading comprehension has been established by second language acquisition (SLA) research (e.g. Nation & Coady, 1988; Laufer, 1992, 1996; Wallace, 2007). Especially the role of vocabulary breadth, i.e., the size of the learner's vocabulary, has been found to be important in second language proficiency in general (Henriksen, 2006; Schmitt, 2010) and in reading competence in particular (Laufer, 1992, 1996; Laufer & Ravenhorst-Kavlovski, 2010). The other dimension of vocabulary knowledge, i.e. depth, or quality, has not been as thoroughly researched, although its overall role in second language vocabulary competence has been acknowledged (e.g., Meara, 1996; Read, 2000). The objective of the study reported in this article was to examine the relationship between the breadth and depth of vocabulary knowledge of Finnish learners of English, representing intermediate and advanced proficiency levels. Additionally, acknowledging the well-established connection between vocabulary size and reading comprehension ability, another aim of the study was to see whether assessing vocabulary depth could add a viable dimension in predicting and explaining reading comprehension proficiency. In the following, the study will first be situated in the relevant framework of earlier research.

II. VOCABULARY BREADTH AND DEPTH

Vocabulary breadth, or size, is usually understood to denote the number of words a language learner knows. Depth of vocabulary knowledge, on the other hand, commonly refers to how well these words are known. However, the terms are far from straightforward. There is no general consensus of what is meant by *knowing a word* or how *depth of vocabulary knowledge* can be defined. Following Nation's (2001) tripartite division of word knowledge into form, meaning, and use, Li and Kirby (2015, p. 612) suggest that breadth of vocabulary could be seen as "knowing the oral and written forms of the words, the surface meanings, and basic uses of the words." Other researchers emphasize the importance of knowing the meaning of the word. Qian (1999, 2002), for example, regards vocabulary size as the number of words for which the learner has at least some superficial knowledge of meaning.

Defining depth of vocabulary knowledge seems to be an even more complex task. Indeed, there are at least two rather different approaches to conceptualising depth of vocabulary knowledge: the developmental approach and the dimensions (or components) approach (Read, 2000; Schmitt, 2010). According to the developmental approach, depth of vocabulary knowledge grows incrementally, from not knowing a word at all, through recognition and having a vague idea of its meaning, to mastering the word, whereas according to the dimensions approach, knowledge of a word is seen as consisting of different subcomponents, or different types of word knowledge, such as orthographic, phonological, morphological, semantic, syntactic, collocational, and pragmatic features (Read, 2004; Schmitt, 2010). Understandably, it is impossible to design tests which would measure all of these features. Some attempts have been made, nevertheless, to include several aspects of vocabulary knowledge in test formats, as reported by Read (2000). The approach adopted in

the present study is the developmental one, i.e. depth of vocabulary knowledge is understood as a sequence of developing stages of word knowledge. Consequently, the test employed to measure the study subjects' vocabulary depth was the Vocabulary Knowledge Scale (VKS), developed by Wesche and Paribakht (1996), as explained in the section on *Methods and Procedures*.

Vocabulary breadth and vocabulary depth have both received considerable attention from researchers, but there is some controversy about these two constructs. In addition to the multifaceted nature of the depth of vocabulary knowledge and the consequent difficulty in measuring it, the mutual relationship of breadth and depth remains a contentious issue. They are certainly related to each other, as high correlations have been found between them (e.g. Qian, 1999, 2002; Nurweni & Read, 1999, especially for high proficiency students). According to Li and Kirby (2015), breadth and depth could be seen as two dimensions of the same phenomenon which are interconnected and influence each other. The same idea has been expressed by other scholars as well: a beginning learner recognizes words and learns their basic meanings. With growing experience, more and more words are known, and their characteristics, use, and associations with other words become more familiar. As Li and Kirby (2015, p. 613) put it, "depth contributes to breadth and vice versa."

III. VOCABULARY KNOWLEDGE AND READING COMPREHENSION

A. Text Coverage and Other Issues

Despite the central role of vocabulary in all four L2 skills (reading, writing, listening, and speaking), there are some differences between the written and spoken modes. It has been suggested that more knowledge of words is required for writing and reading than for speaking and listening (e.g. Nation, 2001; Nation, 2006; Schmitt, 2008). Furthermore, it has been proposed that poor L2 reading performance might result from poor L1 reading skills, but available reading research strongly suggests that L2 reading is more of a language problem than an actual reading problem (Alderson *et al.*, 2015). Especially vocabulary knowledge is vital in reading comprehension, and this has been supported by many studies (e.g. Nation & Coady, 1988; Laufer, 1992). Vocabulary knowledge is actually understood to be the most obviously perceivable component of the reading skill (Nation & Coady, 1988). In addition, previous research undeniably shows that the link between reading comprehension and vocabulary knowledge is stronger than the one between reading comprehension and other subcomponents of reading, such as grammar knowledge or background knowledge (Mehrpour & Rahimi, 2010; Alderson *et al.*, 2015). Laufer (1996) states that lexical problems undeniably hinder reading comprehension and that the threshold for successful reading is primarily related to vocabulary knowledge. In the case of EFL in particular, vocabulary size strongly limits the number of texts a learner can read with ease (Nation & Meara, 2002).

Many studies have examined the vocabulary size needed to understand a variety of texts. Laufer (1992) studied L2 lexical knowledge and the readers' general academic ability and concluded that on all levels of knowledge, L2 vocabulary size is more efficient in predicting L2 reading performance than the informant's general academic ability. She suggests that knowing about 3000 word families is enough to ensure a starting point for L2 reading comprehension. Correspondingly, Nation and Waring (1997) propose that the lexical threshold for reading comprehension is somewhere between 3000 and 5000 word families. Many researchers, however, propose that somewhat larger vocabularies are needed for L2 reading to be successful (e.g. Nation, 2006; Schmitt, 2008). Estimates of *text coverage* needed for successful L2 reading vary considerably: according to many, it is somewhere between 95 and 99% (Mehrpour & Rahimi, 2010). Text coverage needed for reading comprehension refers to how many word families or lexical units are needed for understanding a text; in other words, it is the percentage of running words that a learner needs to know in order to comprehend a given text (Nation, 2006). A 95% text coverage would mean that approximately one word out of twenty is unknown, and with a 98% text coverage one word out of fifty would be unknown. In recent studies, the most commonly suggested text coverage figure is 98% (e.g. Schmitt, Jiang & Grabe, 2011). Similarly, Nation (2006) suggests that a 98% text coverage is needed for satisfactory, unaided reading comprehension. In practice, this means that a learner is required to know between 8000 and 9000 word families. To compare, a well-educated native speaker of English is estimated to have knowledge of about 20000 word families. The number of words needed for reading depends, naturally, also on the type of text (see e.g. Nation, 2006).

As the present study is concerned with the vocabulary knowledge of Finnish learners of English, the results of some prior studies investigating the vocabulary size of this learner group are in order. In an early study (1993), Jaatinen and Mankinen discovered that MA level university students of English knew about 19 500 lexemes. More recently, Pirilä (2012) found that the vocabularies of her informants, also MA level English majors, comprised about 23 200 lexemes. It is worth noting that both of these studies counted lexemes instead of word families, and their method of measuring the receptive vocabulary size of their informants was a Yes/No test, built on dictionary entries. Ala-Akkala (2010) focused on upper secondary school students, i.e. intermediate level English learners, and found their receptive vocabulary breadth to be 3700 word families on average. Her method was also a Yes/No test. The method used in the present study to measure vocabulary breadth had a multiple-choice format. This will also be explained in more detail in the section on *Methods and Procedures*.

B. Breadth, Depth, and Reading

Studies concerning the relationship between vocabulary size, or breadth, and reading comprehension are numerous,

and the connection between the two is well established, as was seen above. The number of studies concerning both vocabulary breadth and depth as well as reading comprehension is more limited, however. High positive correlations between vocabulary breadth, depth, and reading comprehension were found by Rashidi and Khosravi (2010), whose study comprised Iranian EFL learners. A moderate correlation was found between breadth and depth of vocabulary by Li and Kirby (2015), but the relationship of these two constructs to reading comprehension showed intriguing results, as vocabulary breadth correlated more strongly with a multiple-choice task, whereas depth of vocabulary correlated more strongly with a more demanding summary task.

Qian has been one of the pioneers in acknowledging the significance of depth of vocabulary knowledge as a vital element of reading comprehension skills along with vocabulary size. In his 1999 study, high positive correlations were established between the scores obtained in the four tests used to measure vocabulary breadth, vocabulary depth (two different tests), and reading comprehension. Moreover, both vocabulary breadth and depth provided significant contributions to the prediction of reading proficiency. What is noteworthy here is that vocabulary depth indeed added 11 percentage points of explained variance in reading test scores beyond the prediction provided by vocabulary breadth alone. The study undeniably supports the significance of vocabulary depth in reading comprehension (Qian, 1999). Qian's 2002 study lends further support to the importance of vocabulary depth as a predictor of reading comprehension, as vocabulary depth scores alone explained about 59% of the variance of the results, whereas the scores on vocabulary breadth alone explained about 54% of the same variance (Qian, 2002).

IV. THE STUDY

A. The Aim of the Study

The main purpose of the present study was to examine how vocabulary breadth, vocabulary depth, and reading comprehension relate to each other. Another aim was to discover the role of vocabulary depth in the prediction and explanation of L2 learners' reading comprehension. The third aim of the study was to compare two groups of learners of English at different proficiency levels, intermediate and advanced, to see whether there were any differences regarding the first two research questions between the groups.

B. Subjects

The subjects of the study consisted of two groups, 39 upper secondary school students from a medium-sized town in Southern Finland (the UPSEC group) and 19 university students, majoring in English at a large university in South-West Finland (the UNI group). The majority of the subjects were female (67% in the UPSEC group and 89% in the UNI group). 88% of the participants had Finnish as their L1 (82% of the UPSEC group and 100% of the UNI group). As even those subjects who had a different L1 were all attending a school with Finnish as the primary language of instruction, and as they all reported using Finnish constantly in their everyday lives, we decided not to exclude them from the study. Some basic information about the groups of subjects is presented in Table I.

TABLE I.
SUBJECTS OF THE STUDY

	UPSEC (n = 39)	UNI (n = 19)
Gender	F: 26 M: 13	F: 17 M: 2
Age (mean)	17.3	27.0
L1	Finnish: 32 Other: 7	Finnish: 19 Other: 0
Started English at age	9: 30 earlier: 1 later: 8	9: 14 earlier: 2 later: 3
Proficiency level	B1-B2	C1-C2

The background information form that the subjects were asked to fill in also included a question about possible longer stays in an English-speaking country. None of the UPSEC students reported such stays; nine UNI students did report stays of varying lengths. However, as all the subjects in the UNI group were English majors who were already at a very advanced level in their studies, the length or number of stays in an English-speaking environment was not considered a major factor in this study.

C. Methods and Procedures

The subjects' vocabulary breadth was measured using the Vocabulary Size Test (VST), designed by Nation and Beglar (2007), and available at http://www.lex tutor.ca/tests/levels/recognition/1_14k/. The VST is based on frequency levels sampled according to the occurrence of word families in the British National Corpus (BNC). It has a monolingual multiple-choice format where the learner has to circle the alternative that is closest in meaning to the item being defined. Each test item is presented in a "short, non-defining context" (Nation & Beglar, 2007: 12). Furthermore, whenever feasible, the words used in the alternatives are of higher frequency than the test item in question. An example is provided below:

- STONE: He sat on a **stone**.
- a. hard thing
 - b. kind of chair
 - c. soft thing on the floor

d. part of a tree

Scoring the test is very straightforward. For each test item, there is only one correct answer, and each test item is worth one point. The maximum score for the test is, thus, 140 points. This number has to be multiplied by 100 to get an estimate of the informant's total vocabulary size (Nation & Beglar, 2007). For the purposes of this study, mainly because of time constraints, however, the test was shortened. The number of test items was reduced to 100: the last forty questions testing the 11K–14 levels were deleted. According to the authors of the test, this is perfectly appropriate, as long as the informants do a few levels beyond their expected proficiency level (Nation & Beglar, 2007). This way, it was possible to do this part of the test session within the time limit of twenty minutes.

As mentioned earlier, the view adopted in the present study on depth of vocabulary knowledge is based on the developmental approach, i.e. depth of vocabulary knowledge is understood to grow gradually, from no knowledge to mastery. A test designed by Wesche and Paribakht (1996), the Vocabulary Knowledge Scale (VKS), was, therefore, used to measure the subjects' vocabulary depth. The VKS combines both self-report and performance items, ratings varying from utter unfamiliarity ("I don't remember having seen this word before"), through word recognition ("I have seen this word before, but I don't know what it means") and having a vague idea of the meaning ("I have seen this word before, and I *think* it means ___"), to the learner being able to construct a sentence where the word is correctly used (Wesche & Paribakht, 1996, p. 29). The test was scored according to instructions given by the authors (Wesche & Paribakht, 1996; Paribakht & Wesche 1997) and Joe (1998).

As for the words to be included in the VKS test, we chose fourteen words, one from each frequency band of the BNC-COCA frequency list, using the online program VocabProfile (<http://www.lexutor.ca/vp/comp/>): *strong, equipment, environment, interpretation, confident, precious, diary, exhilarating, doughnut, to amputate, infantile, rapacious, wallflower, parsonage*. The words came from texts which were randomly selected from an upper secondary school textbook. As the words represented such a wide range of frequency levels, they were expected to reveal differences in the informants' depth of vocabulary knowledge.

The third test in the present study was a reading comprehension test, which was based on sections of the English test of the Finnish matriculation examination. This examination is designed and created every year by the Finnish Matriculation Examination Board. The examination is taken simultaneously every autumn and spring by Finnish upper secondary school students usually in their third and final year of their studies. In the English test, the reading comprehension section most commonly consists of a number of texts that are accompanied by multiple choice questions. The tests from previous years along with the correct answers and comments from the Matriculation Examination Board can be found online (<http://yle.fi/aihe/artikkeli/2015/12/15/yo-kokeet-englanti>), but in order to use the tests for research purposes, consent has to be acquired from the Board. We applied for a research permission in order to be able to use the reading comprehension tests for the purposes of the present study, and the permission was granted.

We chose two reading comprehension sections, one from autumn 2012 and one from spring 2013, read carefully all texts, questions and responses, and came up with a reading comprehension test that was comprised of three texts accompanied by corresponding multiple choice questions. The texts represented different fields, the first text being an extract from a novel, the second an editor's note from *National Geographic*, and the third an article from a magazine. For each text, there were three to five multiple choice questions, each with three alternatives in English. The UNI students were not likely to be familiar with the tests from such recent years. As for the UPSEC students, it was confirmed by their teachers that these tests had not been used as practice material in class.

The data were analysed using *IBM SPSS Statistics 23.0.0.2*. In order to examine the relationships between different test results and between the performances of the two study groups, the *correlation coefficient* (r) was used. As the results in one of the tests were so uniform that the scores actually behaved more like a categorical variable than a continuous one, and, in addition, the scores were not normally distributed, a non-parametric test, *Spearman's rho*, was applied. As a means of describing the variance in reading comprehension scores explained by vocabulary breadth and depth, we used *multiple linear regression*. This method of analysis makes it possible to examine the relationship between one dependent variable and one or more independent variables. Multiple linear regression also reveals the amount of variance in the dependent variable explained by one or more independent variables, and this is expressed in R square (R^2) values that vary from 0 to 1 (Muijs, 2004).

V. RESULTS

The breadth of vocabulary knowledge (or vocabulary size) of the subjects was measured with the Vocabulary Size Test (VST), as explained above. The maximum score in the test was 100 points. In order to arrive at an estimation of a learner's vocabulary size, the score then had to be multiplied by 100. Table II presents the central findings of the informants' estimated vocabulary sizes, for each informant group separately and also as a total, calculated in word families. All vocabulary measures here were rounded to integers.

TABLE II.
SCORES ON THE VOCABULARY SIZE TEST (VOCABULARY BREADTH)

	Mean	St. Dev.	Min.	Max.	Median
UPSEC (n = 39)	5715	1761	1500	8500	5600
UNI (n = 19)	8816	637	7300	9800	8900
Total (n = 58)	6731	2086	1500	9800	7300

As can be seen in Table II, the UNI group performed better in the VST than the UPSEC group. The mean vocabulary size of the UNI informants was 8816 words, whereas that of the UPSEC group was 5715. A Mann-Whitney U-test was conducted in order to verify the statistical significance of the difference in the results of the two study groups ($U = 23.0, p < 0.01$).

The informants' depth of vocabulary knowledge was assessed with the Vocabulary Knowledge Scale (VKS). The maximum score for this test was 70 points. The results of the two groups are shown in Table III. The measures were again rounded to integers, except for standard deviation, which is presented to one decimal place.

TABLE III.
SCORES ON THE VOCABULARY KNOWLEDGE SCALE (VOCABULARY DEPTH)

	Mean	St. Dev.	Min.	Max.	Median
UPSEC (n = 39)	36	11.3	16	58	36
UNI (n = 19)	60	3.5	52	66	60
Total (n = 58)	44	14.7	16	66	45

In the VKS test, the mean score for the UPSEC group was 36 points out of 70. The least successful participant scored 16 points, whereas the most successful UPSEC student scored 58 points. Similarly to the scores in the VST, the most successful participants in the UPSEC group were close to the mean score of the more advanced UNI group. Again, the standard deviation of the UPSEC group's scores was large: 11.3. In the UNI group, on the other hand, the differences between the participants were less dramatic. The mean score was 60 points and the standard deviation only 3.5. In this group, the least successful informant scored 52 points, compared to the most successful informant's 66 points. A Mann-Whitney U-test indicated that the difference in the means between the two groups was statistically significant ($U = 11.5, p < 0.01$).

The reading comprehension test used in this study comprised of three different texts accompanied by multiple-choice questions. The maximum score in this section was 12 points. The most central findings of this test can be seen in Table IV. Again, the measures were rounded to integers for clarity (except for standard deviation that was rounded to one decimal place).

TABLE IV.
SCORES ON THE READING COMPREHENSION TEST

	Mean	St. Dev.	Min.	Max.	Median
UPSEC (n = 39)	6	2.8	0	11	7
UNI (n = 19)	11	1.1	9	12	11
Total (n = 58)	8	3.2	0	12	8

In the UPSEC group, the differences between the learners were again noteworthy. The least successful informant did not manage to provide any correct answers to the multiple-choice questions, whereas the most successful informant scored 11 out of 12 points. The standard deviation was 2.8. The mean score for this section, in the UPSEC group, was 6 points. Similarly to the scores in the other two tests reported above, the differences between the participants' results in the UNI group were less notable. In this group, the lowest score was 9 points and the highest 12 points. The standard deviation was 1.1. The mean score of the informants in the UNI group was 11 points. As with the other two test results presented above, a Mann-Whitney U-test was conducted, and the difference in the means of the UPSEC and UNI groups was statistically significant ($U = 40.0, p < 0.01$).

To examine the correlations between the test results, the Spearman's rho was used. Table V shows the UPSEC group's correlations between the results of the three tests.

TABLE V.
CORRELATIONS IN THE UPSEC GROUP

	VST (Vocabulary breadth)	VKS (Vocabulary depth)	RC (Reading comprehension)
VST		0.89	0.65
VKS	0.89		0.59
RC	0.65	0.59	

As Table V indicates, the correlations in the UPSEC group were rather high and positive across all test results. Vocabulary size and the scores on the VKS correlated very strongly and positively ($r = 0.89, n = 39, p < 0.001$). Vocabulary size also correlated strongly with the reading comprehension test results ($r = 0.65, n = 39, p < 0.001$). The correlation between the scores on the VKS and on the reading comprehension test was also strong ($r = 0.59, n = 39, p < 0.001$).

In the UNI group, on the other hand, the correlations were not as high as in the UPSEC group (see Table VI).

Similarly to the UPSEC group, vocabulary size and the scores on the VKS did correlate strongly ($r = 0.72$, $n = 19$, $p = 0.001$) in the UNI group as well, but the other two correlations were clearly weaker. There was a moderate positive correlation between vocabulary size and reading comprehension scores ($r = 0.31$, $n = 19$, $p = 0.205$). A modest positive correlation was also found between the scores on the VKS and the reading comprehension test ($r = 0.29$, $n = 19$, $p = 0.224$). As can be seen from the p-values, the correlations between vocabulary size and reading comprehension, as well as between vocabulary depth and reading comprehension scores were not statistically significant ($p > 0.05$).

TABLE VI.
CORRELATIONS IN THE UNI GROUP

	VST (Vocabulary breadth)	VKS (Vocabulary depth)	RC (Reading comprehension)
VST		0.72	0.31
VKS	0.72		0.29
RC	0.31	0.29	

In Table VII below, the correlations for both groups together are presented. The correlations were strong and positive across all variables. A very strong, positive correlation was found between vocabulary breadth (VST) and vocabulary depth (VKS) ($r = 0.95$, $n = 58$, $p < 0.001$). Vocabulary size was also very strongly and positively related to the scores in the reading comprehension test ($r = 0.83$, $n = 58$, $p < 0.001$). Another very strong, positive correlation was found between the scores on the VKS and the reading comprehension test ($r = 0.95$, $n = 58$, $p < 0.001$).

TABLE VII.
CORRELATIONS IN BOTH GROUPS TOGETHER

	VST (Vocabulary breadth)	VKS (Vocabulary depth)	RC (Reading comprehension)
VST		0.95	0.83
VKS	0.95		0.95
RC	0.83	0.95	

Based on the above findings, the differences between the results of the two study groups are evident. These differences will be further scrutinized in the *Discussion*.

One of the aims of the present study was to examine the role of vocabulary depth in reading comprehension. More precisely, we wanted to see whether depth of vocabulary knowledge would bring something more into the prediction and explanation of reading comprehension skills, in addition to the information afforded by vocabulary breadth alone. Multiple linear regression was chosen in order to answer this research question. In spite of some problems connected with this method, the use of multiple linear regression has been encouraged by existing research (e.g. Qian, 1999, 2002; Qian & Schedl, 2004; Rashidi & Khosravi, 2010), even with independent variables that correlate with each other, as is the case in the present study. Contrary to these studies, however, we decided to report adjusted R^2 values instead of mere R^2 values. R^2 refers to the amount of variance in a dependent variable (in this case, reading comprehension) that is explained by independent variables (vocabulary breadth and depth). An adjusted R^2 is a correction to R^2 that “takes into account that we are looking at a sample rather than at the population” (Muijs, 2004, p. 165). Table VIII shows the results of the regression analysis of the UPSEC group.

TABLE VIII.
MULTIPLE REGRESSION RESULTS OF THE UPSEC GROUP WITH VOCABULARY SIZE AND THE SCORES ON THE VKS AS INDEPENDENT VARIABLES

Step	Variable	Adjusted R^2	R^2 change	p-value
1	Vocabulary size	0.38		<0.001
2	Scores on the VKS	0.39	0.01	<0.001

As summarized in Table VIII, for the UPSEC group, vocabulary size alone explained about 38% of the variance in reading comprehension scores (adjusted $R^2 = 0.38$, $F(1, 37) = 24.68$, $p < 0.001$). When looking at the explanation afforded by vocabulary size and the scores on the VKS together, the percentage was 39% (adjusted $R^2 = 0.39$, $F(2, 36) = 12.88$, $p < 0.001$). Adjusted R^2 change was 0.01, which suggests that vocabulary depth added 1 percentage point of explained variance in reading comprehension scores. When looking at the explained variance provided by vocabulary depth alone, the percentage was about 38% (adjusted $R^2 = 0.38$, $F(1, 37) = 23.77$, $p < 0.001$).

The multiple regression analysis results in the UNI group were rather different from the UPSEC group. Vocabulary size explained only about 19% of the variance in reading comprehension test scores (adjusted $R^2 = 0.19$, $F(1, 17) = 5.32$, $p < 0.05$). Furthermore, when scores on the VKS were added to the equation, it looks as if the two independent variables together are not very good at predicting or explaining reading comprehension test scores: the percentage was only 15% (adjusted $R^2 = 0.15$, $F(2, 16) = 2.62$, $p > 0.05$). The predictive power of vocabulary depth alone was less than one percent (adjusted $R^2 = 0.03$, $F(1, 17) = 1.06$, $p < 0.05$). As can be seen in the p-values reported, only the variance explained by vocabulary size alone is statistically significant. The results seem to suggest, firstly, that neither vocabulary size nor vocabulary depth are very good at predicting or explaining reading comprehension proficiency of the UNI group, and secondly, that scores on the VKS do not actually add anything into the prediction of reading comprehension in this group. A summary of the findings is presented in Table IX.

TABLE IX.
MULTIPLE REGRESSION RESULTS OF THE UNI GROUP WITH VOCABULARY SIZE AND THE SCORES ON THE VKS AS INDEPENDENT VARIABLES

Step	Variable	Adjusted R ²	R ² change	p-value
1	Vocabulary size	0.19		<0.05
2	Scores on the VKS	0.15	-0.04	>0.05

Lastly, we looked at the two groups together to see what the effects of vocabulary size and vocabulary depth were in reading comprehension in general. Vocabulary size alone explained about 64% of the variance in reading comprehension test results (adjusted R² = 0.64, F (1, 56) = 102.17, p<0.001). When both vocabulary size and depth of vocabulary knowledge scores were added to the equation, they together explained about 66% of the variance in reading comprehension (adjusted R² = 0.66, F(2, 55) = 55.07, p<0.001). A modest addition, about 2 percentage points, was afforded by scores on the VKS. When looking at the percentage of explained variance in reading comprehension afforded by vocabulary depth alone, the figure was 64% (adjusted R² = 0.64, F (1, 56) = 103.95, p<0.001). Table X below sums up the findings.

TABLE X.
MULTIPLE REGRESSION RESULTS OF BOTH GROUPS TOGETHER WITH VOCABULARY SIZE AND THE SCORES ON THE VKS AS INDEPENDENT VARIABLES

Step	Variable	Adjusted R ²	R ² change	p-value
1	Vocabulary size	0.64		<0.001
2	Scores on the VKS	0.66	0.02	<0.001

As with the correlation coefficients reported earlier, the differences found between the two study groups are again obvious. The findings, and possible explanations for them, will be further discussed in the following chapter.

VI. DISCUSSION

The primary objective of the study at hand was to investigate how L2 vocabulary breadth, vocabulary depth, and reading comprehension relate to one another, specifically focusing on the role of depth of vocabulary knowledge in predicting and explaining reading comprehension skills. A further goal was to discover whether the results would be different in two proficiency groups, intermediate and advanced.

The receptive vocabulary size of the group of upper secondary school students (UPSEC) in the present study was 5715 word families on average. As mentioned earlier, Ala-Akkala (2010) also studied Finnish upper secondary school students and found that their average English vocabulary size was 3700 word families. The method used in her study was a Yes/No test, so it is not directly comparable to the present study, as the difference in testing methods may well have influenced the results. However, the results of both studies seem to suggest that the differences between individual upper secondary school students are large.

As for the group of university students (UNI), on average, a major student of English knew 8816 word families. However, it is noteworthy to remember that the version of the VST used in the present study only tested the knowledge of the 10 000 most frequent word families of English. In order to tap the whole scope of the learners' vocabulary breadth, another test involving further frequency levels as well ought to be used. It is likely that the results obtained would have been somewhat higher had the test included words from the lowest frequency levels as well.

A major difference between earlier Finnish studies on learners' vocabulary sizes and the present study lies in the test format. As pointed out before, the Yes/No and the multiple choice test formats are quite different. The Yes/No test format does not require the informant to actually show any knowledge of word meaning; the method is more like self-assessing one's knowledge. In the multiple-choice format, which was applied in the present study, however, the informant is required to demonstrate his or her knowledge. On the other hand, the multiple-choice format mostly requires the informant to know one particular meaning of the test item in question, whereas in the Yes/No format, the informant can report any meaning out of all the meanings of a particular word. Thus, it is likely that the use of multiple-choice formats results in somewhat smaller estimates of learners' vocabulary size than the use of tests based on the Yes/No format. In addition, as mentioned earlier, with the VST applied in this study, the maximum vocabulary size that the test could report was 10000 word families. A longer test or a test compiled differently (for example, five test items instead of ten to represent one frequency level) might be useful in order to better estimate the vocabulary size of advanced learners of English. Even though the authors of the VST, Nation and Beglar (2007), accept shortening the test, they suggest, nonetheless, that the best is to strive for a test where the last frequency levels to be tested are a few levels beyond the learners' expected vocabulary size. This may explain why the results of the UNI group differ so drastically from other studies conducted with Finnish university students of English (Jaatinen & Mankkinen, 1993; Pirilä, 2012).

Positive correlations were found between L2 vocabulary size, depth and reading comprehension. The correlations were especially high in the UPSEC group and in the two groups together. In the UNI group, on the other hand, the correlations were somewhat lower. Many researchers have pointed out that the relationship between the size and depth of vocabulary knowledge is a close one (e.g. Qian, 1999). The findings of the present study support this view. The correlation between vocabulary size and the scores on the VKS was high and positive in both groups: 0.89 in the UPSEC group, 0.72 in the UNI group, and 0.95 in the two groups together. Similar results have been reported in earlier studies as well (Rashidi & Khosravi, 2010; Li, 2015).

In this study, the correlation between vocabulary size and reading comprehension was positive in both groups. The results are similar to those obtained in earlier studies (e.g. Qian, 1999, 2002; Rashidi & Khosravi, 2010; Li, 2015). Vocabulary depth and reading comprehension were also positively correlated in the present study. This is also in line with results reported in previous research. Rashidi and Khosravi (2010) reported a very strong and positive correlation between vocabulary depth and reading comprehension, whereas Li (2015) reported a slightly lower, yet positive correlation between the two variables. In both of Qian's studies (1999, 2002), two different measures were used to assess the learners' depth of vocabulary knowledge. Nonetheless, the correlations still resemble the ones already mentioned here, as do the results obtained in a study by Qian and Schedl (2004), which also focused on the relationship between vocabulary depth and reading comprehension.

As for the second research question, the results of the present study suggest that vocabulary depth can indeed add another significant variable in the prediction and explanation of scores in a reading comprehension test. In the UPSEC group, vocabulary breadth alone explained 38% of the variance in reading comprehension. Together with vocabulary depth, the two variables explained 39% of the variance. When examining the UPSEC and UNI groups together, the corresponding figures were 64% and 66%, respectively. These results suggest that vocabulary depth added 1 percentage point and 2 percentage points of explained variance in reading comprehension scores when considering the UPSEC group and the UPSEC and UNI groups together. Similarly to the figures of variance explained by vocabulary size alone, vocabulary depth alone could explain 38% and 64% of variance in reading comprehension scores in the UPSEC group and in the two groups in total, respectively. It seems that vocabulary depth does not add much to the explanation and prediction of L2 reading, in addition to that afforded by vocabulary size alone. Nevertheless, the results suggest that both vocabulary size and depth could be used separately as equivalent predictors of reading proficiency. However, when looking at the results of the UNI group, the results were somewhat different. In the UNI group, vocabulary size alone explained 19% of the variance in reading comprehension. Together with vocabulary depth, the two variables explained only 15% of the variance. Within the scope of the present study, the numbers suggest that vocabulary depth does not necessarily provide any additional information which would not be provided by vocabulary size alone, in the prediction of reading comprehension skills of the highly advanced UNI group. Moreover, when looking at the predictive and explanatory power of vocabulary depth alone, it could only explain less than one percent of the variance in reading comprehension test scores.

In their study, Rashidi and Khosravi (2010) only looked at the separate contributions afforded by vocabulary breadth and depth in reading comprehension. According to them, vocabulary size accounted for 55% and vocabulary depth for 69% of the variance in reading comprehension test scores. The results are rather similar to those obtained in the present study where vocabulary size alone explained 64%, and, similarly, vocabulary depth alone explained 64% of variance in reading comprehension test results, when looking at the two groups of informants in total. In Qian and Schedl's (2004) study, only vocabulary depth and reading comprehension were involved. The scores of their vocabulary depth test explained 55% of the variance in reading comprehension test scores.

In Qian's 1999 study, vocabulary size alone explained 60% of the variance in reading comprehension scores. Together with vocabulary depth, the two variables could explain 71% of the variance. This suggests that vocabulary depth added significant 11 percentage points of explained variance in reading comprehension test results.

When examining the results of earlier studies and the present study, it seems that measuring vocabulary depth can indeed serve as a tool for predicting and explaining reading comprehension test scores. Both vocabulary size and depth alone were able to explain a considerable and similar amount of variance in reading comprehension test scores. However, the percentage of added explained variance provided by vocabulary depth (in addition to vocabulary size alone) was not very high, and, in the present study, it was indeed lower than in previous studies (Qian, 1999, 2002). It might be that the test format played a role here because the VKS used in the present study is fundamentally different from the tests used in the majority of earlier research to measure vocabulary depth (as mentioned earlier, the VKS reflects the developmental approach to vocabulary depth, whereas most other tests are based on the dimensional view). Nevertheless, the results of the present study suggest that measuring vocabulary depth as a means of predicting or explaining reading comprehension test scores is quite as efficient as measuring vocabulary size.

As reported and discussed already, the UNI group performed better than the UPSEC group in all three tests. This is not a surprising finding, considering that the two groups were of different competence levels, the UPSEC group consisting of intermediate-level language learners and the UNI group of advanced learners. Moreover, the differences within the two groups varied considerably. The more advanced UNI group was clearly quite homogeneous, whereas the differences within the less advanced UPSEC group were large. It has to be acknowledged, of course, that the UNI group was somewhat smaller than the UPSEC group, which may also have affected the results.

When it comes to the correlations found between the three variables (vocabulary size, vocabulary depth, and reading comprehension), they were stronger in the UPSEC group than in the more advanced UNI group. This seems to suggest that the informants in the UNI group performed well in the reading comprehension test despite the lower correlations between vocabulary size and reading comprehension, as well as between vocabulary depth and reading comprehension. Considering their very high competence of English and their level of studies, it is likely that they were able to use more effective reading strategies, for example. In addition, the tests conducted may simply have been at least partly too easy for the UNI group. The significant finding here is that the results suggest that vocabulary knowledge, understood both

as vocabulary size and vocabulary depth, plays a fundamental role in reading comprehension proficiency of learners who are not very advanced, but the role of vocabulary knowledge seems to decrease with increasing L2 proficiency.

VII. CONCLUSION

Vocabulary knowledge has a central role in L2 learning in general and in reading comprehension specifically. The results of this study strongly suggest that vocabulary breadth, depth and reading comprehension are profoundly connected with each other, as has, indeed, been suggested by many previous research findings. This is especially interesting from the viewpoint of Finnish EFL learners. Finnish is not a cognate of English, nor of any of the languages that were any of the informants' L1 mentioned in the previous studies of the field (e.g. Qian, 1999, 2002; Qian & Schedl, 2004; Rashidi & Khosravi, 2010; Li, 2015). The results of the present study, as well as the earlier ones, support the view that L2 reading is fundamentally a language problem, which cannot be explained only by the learner's L1 or L1-related skills, and L2 vocabulary knowledge strongly affects reading comprehension.

The results of the present study suggest that both vocabulary breadth and depth can be used as predictors of reading comprehension proficiency, as they seem to provide a similar factor in the explanation of reading comprehension test scores. However, caution has to be applied when considering advanced learners. It seems that at the more advanced proficiency levels, the learners will succeed in reading comprehension despite their vocabulary knowledge. It is likely that advanced learners have a wide range of other skills, such as the use of other linguistic knowledge or reading strategies, which they can resort to when encountering problems in reading. At the beginning and developing stages of competence, on the other hand, both vocabulary size and depth are good predictive and explanative factors of reading comprehension.

ACKNOWLEDGEMENT

The authors wish to thank the anonymous reviewers for their observant and insightful comments.

REFERENCES

- [1] Ala-Akkala, M. (2010). The development of English vocabulary during upper secondary school. Unpublished MA thesis, University of Turku.
- [2] Alderson, C. J., E. Haapakangas, A. Huhta, L. Nieminen & R. Ullakonoja. (2015). The diagnosis of reading in a second or foreign language. New York: Routledge.
- [3] Henriksen, B. (2006). Exploring the quality of lexical knowledge in the language learners' L1 and L2. In P. Pietilä, P. Lintunen & H. Järvinen (eds.), *Kielenoppija tänään – language learners of today*. Jyväskylä: AFinLA, 13–36.
- [4] Henriksen, B. & L. Danelund. (2015). Studies of Danish L2 learners' vocabulary knowledge and the lexical richness of their written production in English. In P. Pietilä, K. Doró & R. Pápalová (eds.), *Lexical issues in L2 writing*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- [5] Jaatinen, S. & T. Mankkinen. (1993). The size of vocabulary of university students of English. In K. Sajavaara & S. Takala (eds.), *Finns as learners of English: three studies*. Jyväskylä: University of Jyväskylä, 147–211.
- [6] Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics* 19.3, 357–377.
- [7] Laufer, B. (1992). Reading in a foreign language: How does L2 lexical knowledge interact with the reader's general academic ability? *Journal of Research in Reading* 15.2, 95–103.
- [8] Laufer, B. (1996). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady & T. Huckin (eds.), *Second language vocabulary acquisition: a rationale for pedagogy*. Cambridge: Cambridge University Press, 20–34.
- [9] Laufer, B. & G. Ravenhorst-Kavlovski (2010). Lexical threshold revisited: lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language* 22.1, 15–30.
- [10] Li, K. (2015). A study of vocabulary knowledge and reading comprehension of EFL Chinese learners. *Studies in Literature and Language* 10.1, 33–40.
- [11] Li, M. & J. R. Kirby (2015). The effects of vocabulary breadth and depth on English reading. *Applied Linguistics* 36.5, 611–634.
- [12] Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer & J. Williams (eds.), *Performance and competence in second language acquisition*. Cambridge: Cambridge University Press, 35–53.
- [13] Mehrpour, S. & M. Rahimi. (2010). The impact of general and specific vocabulary knowledge on reading and listening comprehension: a case of Iranian EFL learners. *System* 38.2, 292–300.
- [14] Muijs, D. (2004). *Doing quantitative research in education with SPSS*. London: SAGE Publications Ltd.
- [15] Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- [16] Nation, P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review* 63.1, 59–82.
- [17] Nation, P. & J. Beglar. (2007). A vocabulary size test. *The Language Teacher* 31.7, 9–13.
- [18] Nation, P. & J. Coady. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (eds.), *Vocabulary and language teaching*. London: Longman, 97–110.
- [19] Nation, P. & P. Meara. (2002). Vocabulary. In N. Schmitt (ed.), *An introduction to applied linguistics*. London: Edward Arnold: 35–54.

- [20] Nation, P. & R. Waring. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (eds.), *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press, 6-19.
- [21] Nurweni, A. & J. Read. (1999). The English vocabulary knowledge of Indonesian university students. *English for Specific Purposes* 18.2, 161-175.
- [22] Paribakht, S. T. & M. Wesche. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (eds.), *Second language vocabulary acquisition: a rationale for pedagogy*. Cambridge: Cambridge University Press, 174–200.
- [23] Piriälä, A. (2012). Knowledge of English vocabulary at advanced levels: vocabulary size and academic vocabulary knowledge of Finnish 1st year university students and MA students of English. Unpublished MA thesis, University of Turku.
- [24] Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *The Canadian Modern Language Review* 56.2, 282–307.
- [25] Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning* 52.3, 513–536.
- [26] Qian, D. & M. Schedl. (2004). Evaluation of an in-depth vocabulary knowledge measure for assessing reading performance. *Language Testing* 21.1, 28–52.
- [27] Rashidi, N. & N. Khosravi. (2010). Assessing the role of depth and breadth of vocabulary knowledge in reading comprehension of Iranian EFL learners. *Journal of Pan-Pacific Association of Applied Linguistics* 14.1, 81-108.
- [28] Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- [29] Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (eds.), *Vocabulary in a second language: selection, acquisition, and testing*. Amsterdam: John Benjamins, 209–227.
- [30] Schmitt, N. (2008). Review article – Instructed second language vocabulary learning. *Language Teaching Research* 12.3, 329–363.
- [31] Schmitt, N. (2010). *Researching vocabulary: a vocabulary research manual*. London: Palgrave Macmillan.
- [32] Schmitt, N., X. Jiang & W. Grabe. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal* 95, 26–43.
- [33] Wallace, C. (2007). Vocabulary: The key to teaching English language learners to read. *Reading Improvement* 44.4, 189–193.
- [34] Wesche, M. & S. T. Paribakht. (1996). Assessing second language vocabulary knowledge: depth versus breadth. *Canadian Modern Language Review* 53.1, 13–40.

Noora Harkio is a teacher of English, French and Swedish. Her MA degree, with English as her major subject, is from the University of Turku (2016). She has formerly worked at Sesevetska Sopenica Elementary School in Zagreb, Croatia, and, as of August 2016, will start teaching English and Swedish in an upper secondary school in the capital area of Finland. Her research interests include second language learning, L2 vocabulary acquisition as well as specific learning differences in SLA.

Päivi Pietilä is Professor of English at the University of Turku, Finland, where she is responsible for the SLA section of her department. Her publications include *The English of Finnish Americans* (1989), *L2 Speech* (1999), *Lexical Issues in L2 Writing* (2015, co-editor), in addition to a number of journal articles. Her research interests include second language acquisition and attrition, vocabulary acquisition and use, L2 academic writing, L2 speaking skills, and the lexis-grammar interface.