



VOCABULARY SIZE AT FOUR STAGES OF LANGUAGE DEVELOPMENT

Nina Daskalovskaⁱ

Goce Delcev University, Stip,
Republic of North Macedonia

Abstract:

One of the main aims of language learning is developing communicative competence, or the ability to communicate effectively in everyday situations. In order to achieve that, besides grammar knowledge, learners need to acquire enough vocabulary knowledge that would enable them to use the language outside the classroom, which requires acquiring significant receptive and productive word knowledge in order to be able to participate in various communicative situations. Research shows that the most frequent 2000 words comprise about 85% of the words in any text regardless of the topic, and that learners need to know at least 3000 word families in order to be able to use the language successfully and to be able to read authentic texts with understanding. However, experience shows that acquiring a large vocabulary is one of the biggest obstacles that language learners face in the process of language acquisition. Nevertheless, since vocabulary knowledge is the key component of the overall communicative competence, is it necessary to invest time and effort to learn the most frequent words in the language. The aim of this study is to determine the vocabulary size of language learners at four stages of language development. In addition, the study attempts to determine the difference between receptive and productive knowledge of vocabulary, as well as the difference between knowledge of words seen in isolation and in context. The participants were four groups of learners, of whom three groups comprised elementary and secondary school students and one group were university students. The results showed that during the first five years of learning the vocabulary knowledge increases at a greater rate, that learners have better receptive than productive knowledge of words, and that seeing words in context helps learners demonstrate greater receptive and productive vocabulary knowledge.

Keywords: vocabulary size, receptive knowledge, productive knowledge, isolation, context

ⁱ Correspondence: email nina.daskalovska@ugd.edu.mk, nina.daskalovska@yahoo.com

1. Introduction

Developing the language skills – listening, speaking, reading and writing, is tightly connected to the acquisition of the language systems – grammar, vocabulary and pronunciation. While grammar has always had a central place in language learning programmes, the approach to vocabulary learning until the 1980s was not systematic and it did not have an important role in determining the goals of language learning. Vocabulary started to gain importance with the development of communicative language teaching approaches, especially after the 1980s. Since then, many linguists have emphasized its vital role in language learning and language use. Stressing the importance of vocabulary, Wilkins (1972, p. 111) states that “*while without grammar very little can be conveyed, without vocabulary nothing can be conveyed*”. In the same line, Harmer (1991, p. 153) remarks that “*if language structures make up the skeleton of language, then it is vocabulary that provides the vital organs and the flesh*”. That is why Meara (2002) says that this is a period of rediscovering of vocabulary because we are now asking questions about vocabulary that should have been asked a long time ago and that there is a serious lack of knowledge about the mental lexicon of language learners.

Discussions about vocabulary learning raise a lot of questions: What do we understand by the term vocabulary? What does learning vocabulary involve? How much vocabulary is required to be able to use the language in real-life situations? When we talk about vocabulary, we usually think about single words. However, in addition to single words, the term vocabulary also includes set phrases, variable phrases, phrasal words and idioms (Folse, 2004). Therefore, besides single words, learners need to learn a lot of combinations of words as well. Moreover, learning words does not mean learning only their form and meaning. It requires a multidimensional approach as learners need to acquire many other aspects besides form and meaning. Many linguists have tried to define what it means to know a word. Nation (2001, p. 27) offers a clear and detailed picture of the aspects involved in knowing a word. His classification consists of three main dimensions: form, meaning and use. Each dimension is further divided into three components. Thus, ‘form’ includes written and spoken form and word parts; ‘meaning’ includes form and meaning, concept and referents, and associations. And the last dimension, ‘use’, includes grammatical functions, collocations, and constraints on use, such as register, frequency, etc. And finally, each component is divided into two parts, one of which refers to receptive knowledge, and the other one refers to productive knowledge of words.

Not only do learners need to learn a lot of aspects of word knowledge in order to use them successfully, they also need to learn a large number of words in order to be able to use the language in various situations. According to Meara (1996), the basic dimension of lexical competence is vocabulary size. Learners who have large vocabulary are more successful in all language skills than the learners with small vocabulary. However, Meara states that there are very few studies that measure learners’ vocabulary size, and they include a small number of learners, which makes it difficult to draw firm conclusions. He thinks that the main problem is that there are no adequate tests for measuring vocabulary

size that could shed more light on this issue. One of the tests that is closest to a standard test for measuring vocabulary size is the Vocabulary Levels Test created by Nation (1990) which tests a small number of words grouped according to their frequency using a complex multiple choice format. The first problem that arises is the difficulty in determining the number of words that comprise vocabulary because of the lack of agreement on what can be counted as a word. As a result, the estimations about the vocabulary size of native speakers ranges between 15.000 and 216.00 (Goulden et al., 1990, p. 342). The second problem is that it is very difficult to determine the size of something big on the basis of a small sample (Meara, 1996, p. 39). If someone's vocabulary size is small, it is relatively easy to test a reasonable proportion of the words. For example, if someone knows 400 words, a test with 40 words would be quite representative of the vocabulary size of that person. However, the bigger the vocabulary size, the more difficult it is to provide a reliable test. And the third problem refers to the tests results which greatly depend on the test type. For example, multiple choice tests do not test only the target word, but all the distractors as well. As Meara states, all of these problems have discouraged many linguists to create a standard test for measuring vocabulary size.

2. Receptive versus Productive Word Knowledge

Almost all linguists agree that it is difficult to explain vocabulary development and that despite the numerous studies we still know very little about the processes of vocabulary acquisition. One of the widely accepted views is that it happens along a continuum of development, which means that as learners deepen their knowledge of a word, they move up that continuum. According to some linguists, receptive and productive vocabulary are at the two ends of the continuum. From this perspective, receptive knowledge comes before productive (Waring, 2002).

Melka (1997, p. 84) believes that even though many linguists insist on the dichotomy between reception and production in relation to the lexicon, it is not possible to provide a clear and precise definition of reception and production. Haycraft (as cited in Hatch & Brown, 1995, p. 370) defines receptive vocabulary as the words that learners can recognize and understand when they see them in context, but cannot use them correctly, while productive vocabulary includes the words that learners understand, can pronounce them correctly and can use them constructively in written and oral form. Hedge and Brown (1995, pp. 370-1) also agree that there is not a real dichotomy between receptive and productive vocabulary, but there is a continuum of word knowledge. On the other hand, Meara (1997, pp. 117-8) argues that the metaphors that describe vocabulary acquisition as a progression along a continuum whose end points are defined as receptive/passive and productive/active knowledge, may help us understand vocabulary acquisition, but they have many flaws. He points out that a continuum implies one dimensional space where certain characteristics move in a linear and systematic manner, which cannot be true for the passive/active continuum because passive vocabulary is quite different from active vocabulary (1997, p. 150). Namely, active vocabulary is the vocabulary that is easily accessible from any part of the mental lexicon

and that provides access to other parts of the lexical system through associations, while passive vocabulary consists of words which are part of the system, but they cannot be accessed from other parts of the system. They can only be activated by an external stimulus.

The different estimations of the vocabulary size of native English speakers ranges between 3.000 and 216.000 (Goulden et al., 1990, p. 342). These estimations are based either on word frequency lists or dictionaries. However, the disagreement about the criteria for determining the number of words in the dictionary, as well as the methods for selecting the words has led to quite different estimations. Goulden et al. (1990) conducted a study with 20 native speakers with higher education at the age of 22 in which they attempted to make an objective estimation of the words in the dictionary and their selection. The results showed that adult native English speakers possess a vocabulary of around 17.000 words (p. 354).

Regarding the difference between receptive and productive vocabulary, there are different opinions. According to some estimations, receptive vocabulary is twice the size of productive vocabulary. Eringa (as cited in Melka, 1997, p. 92) estimated that after six years of learning French, high school students have a receptive vocabulary of 4.000-5.000 words and a productive vocabulary of 1.500-2.000 words, while Chamberlain discovered that the receptive vocabulary of an average English speaker is five times larger (10.000 word) than the productive vocabulary (2.000 words). On the other hand, some researchers believe that 92% of receptive vocabulary is known productively and that the figure may be even higher (Schmitt, 2000, p. 119). The differences in the estimations may be due to several reasons, such as the types of words that are tested (frequent/rare words), the criteria and the way the results of the tests are evaluated, the role of context, the role of the native language, the role of cognates, etc. (Melka, 1997, pp. 93-96).

Despite the numerous discussions and studies on the differences between receptive and productive vocabulary, very few studies compare receptive and productive vocabulary learning. According to Waring (1997, p. 99), in order to show receptive knowledge, learners have to provide a translation of the English word into their mother tongue, and to demonstrate productive knowledge, they have to provide the English equivalent of the word in their native language. Stoddard (1929) conducted one of the first studies that investigated receptive and productive vocabulary learning. It included 328 participants, half of whom learned 50 French-English word pairs and the other half learned the same words as English-French pairs. After that, both groups took the same test which tested half of the words receptively (translation from French into English), and half of the words productively (translation from English into French). The results showed that receptive tests are easier than productive tests and that the type of learning affects the test results, i.e. the learners who learned the words receptively had higher scores on the receptive test and vice versa. Waring (1997) conducted a similar test, but the same learners took both the receptive and the productive test. He also investigated the retention of words over time. The results were similar to Stoddard's study. Waring concluded that receptive learning required less time, that the results of the productive test were constantly lower than the receptive test, and that after three months the results

of the productive test were very low. On the other hand, Webb's study (2005) which included two groups of participants showed that the participants who underwent a receptive treatment (reading) had better results on all ten tests which consisted of five receptive and five productive tests than the group who underwent a productive treatment (writing). He concluded that if learners spend equal time on receptive and productive tasks, the receptive tasks, or a combination of both receptive and productive tasks, may be superior to productive tasks not only for developing receptive knowledge, but for gaining productive knowledge as well.

The aim of the present study is to answer the following questions:

- 1) What is the vocabulary size of English language learners at four stages of language development?
- 2) What is the difference between receptive and productive word knowledge?
- 3) Is there a difference in word knowledge if the words are seen in isolation or in context?

3. Material and Methods

3.1. Participants

The participants in this study were 184 Macedonian learners of English at four stages of language development. Therefore, there were four groups of participants: in the first group there were 45 participants, who were twelve-year-old learners who had studied English for two years. In the second group there were 46 fifteen-year-old learners who had studied English for five years. The third group comprised 48 participants who were eighteen years old and had studied English for eight years, and in the fourth group there were 45 participants who were twenty-one years old and had studied English for 11 years. The first three groups of learners were elementary and secondary school students, while the last group were university students of Economics who were following business English courses as compulsory subjects. The first three groups had three class periods a week, while the last group had two class periods a week. The first three groups had a similar composition, there were learners of different language abilities. The last group of participants, who were university students, were randomly chosen from the students enrolled in the third year at the Faculty of Economics.

3.2. Instruments

In order to measure the vocabulary knowledge of the participants, four tests were used. Each test contained 100 words which were nouns, verbs or adjectives, twenty words from each frequency level from 1000 to 5000 words. The first test measured receptive knowledge of words in isolation – the participants were required to provide a synonym or translation of an English word. The second test contained the same words as in the first test, but this time the words were given in a sentence context. For example, the first word in the first test was *accept*, and in the second test the sentence that was used as a context for the word was, *She's decided not to accept the job*. The third and the fourth test contained another set of 100 words, and they tested productive knowledge of words. In

the third test, the participants had to provide the English equivalent of Macedonian words, and in the fourth test, the participants were given an English sentence in which there was a gap which had to be filled with the target word. The target word was given in Macedonian. For example, the English equivalent for the first word in the third test was (рекламира) *advertise*, and the context sentence in the fourth test was: (рекламира) *They no longer _____ alcohol or cigarettes at sporting events*. Thus, the four tests measured both receptive and productive knowledge in isolation and in context. The words were taken from Nation's Vocabulary List (n.d.a), which provides lists of the first ten thousand headwords by frequency levels.

3.3. Procedure

The participants were informed about the aim of the study, and they all agreed to participate. The first two tests were administered on the same day, one after another. First, the participants completed the test where the words were given in isolation, and after that the test with the words in context. The other two tests were administered three days later in the same way. Since there were four tests with 100 words each, it was considered that it would be better to administer them on two different days in order to avoid weariness and boredom.

4. Results and Discussion

Tables 1, 2, 3 and 4 show the results of the four tests. Each frequency band contained 20 words, so both the numbers of the known words and the percentages are given in the tables. The total number of words for the whole test was 100, therefore the total number shows both the number of words the participants knew and the percentage of known words for the group.

Table 1: Receptive knowledge of words in isolation

Test 1											
Participants	1000	%	2000	%	3000	%	4000	%	5000	%	Total
Group 1 (n=45)	5.34	26.72	3.55	17.77	2.07	10.38	1.3	6.5	0.51	2.55	12.77
Group 2 (n=46)	13.45	67.28	9.44	47.22	7.22	36.14	4.47	22.39	3.04	15.21	37.62
Group 3 (n=48)	14.88	74.42	12.11	60.15	9.42	47.13	6.45	32.39	5.5	27.5	48.36
Group 4 (n=45)	14.94	74.16	10.45	51.83	7.67	38.38	5.05	26.72	3.02	15.11	41.13

Table 2: Receptive knowledge of words in context

Test 2											
Participants	1000	%	2000	%	3000	%	4000	%	5000	%	Total
Group 1 (n=45)	6.85	34.27	5.02	25.11	2.78	13.94	2.07	10.38	1.03	5.16	17.75
Group 2 (n=46)	15.27	76.08	11.63	57.5	9.47	47.39	6.26	31.29	5.75	28.53	48.38
Group 3 (n=48)	16.41	82.08	14.37	70.93	11.11	55.57	8.45	42.29	8.44	42.23	58.78
Group 4 (n=45)	17.6	85.66	13.22	66.11	9.76	48.83	7.57	38	5.9	30.34	54.05

Table 3: Productive knowledge of words in isolation

Test 3											
Participants	1000	%	2000	%	3000	%	4000	%	5000	%	Total
Group 1 (n=45)	2.74	13.72	1.76	9.05	2.93	14.66	1.08	5.44	1.85	9.27	10.36
Group 2 (n=46)	9.72	48.64	7.15	35.76	7.78	38.91	3.77	18.85	4.77	23.85	33.19
Group 3 (n=48)	12.93	64.68	8.95	44.79	8.32	41.61	4.15	20.83	5.68	28.43	40.03
Group 4 (n=45)	10.85	54.55	5.51	27.55	4.8	24	2.9	14.5	3.83	19.16	27.89

Table 3: Productive knowledge of words in context

Test 4											
Participants	1000	%	2000	%	3000	%	4000	%	5000	%	Total
Group 1 (n=45)	3.64	17.77	2.52	12.61	3.57	17.88	1.46	7.33	2.05	10.27	13.24
Group 2 (n=46)	11.54	57.93	8.32	47.01	8.7	43.53	4.15	20.8	5.15	25.76	37.86
Group 3 (n=48)	14.15	70.57	10.47	51.25	9.61	48.25	4.97	24.9	6.32	31.61	45.52
Group 4 (n=45)	12.65	60.61	6.66	33.88	5.95	29.66	3.44	17.3	4.44	22.22	33.14

If we accept Nation's (n.d.b) argument that a representative sample of words at one level represents the proportion of all the words known at that level, we will arrive at the numbers given in Table 5, which show the total number of words each group of learners know.

Table 5: The total number of known words

Participants	Test 1 (receptive, in isolation)	Test 2 (receptive, in context)	Test 3 (productive, in isolation)	Test 4 (productive, in context)
Group 1 (n=45)	639	887	519	663
Group 2 (n=46)	1882	2419	1660	1894
Group 3 (n=48)	2418	2940	2003	2277
Group 4 (n=45)	2057	2703	1394	1657

In relation to the first question, we can see that the vocabulary size of English language learners in the first three groups constantly increases, but it does not increase proportionally with the number of years they learn the language. Namely, between the age of twelve and fifteen, both receptive and productive knowledge almost triple, while during the next three years the increase is much slower. So, at the age of eighteen, when they complete their secondary education, on average they know about 2400 words receptively, and about 2000 words productively. The results of the last group of learners show that the knowledge of words at the age of twenty-one has decreased compared to the third group. As this group consisted of university students attending business English courses with a frequency of two class periods per week and the textbooks focused on specialized vocabulary in this area, it is probable that the learners were not exposed to a wide vocabulary like the previous three groups, and they have forgotten some of the vocabulary they had acquired earlier because they did not use it so often during their studies. Moreover, since the words from the Academic Word List were not included, the results of this group of learners may be lower than their real vocabulary knowledge, as

we may assume that English learners at university level may know more words from this list than younger learners.

The second question refers to the difference between receptive and productive knowledge of words. As the figures show, all groups of learners have greater receptive than productive knowledge of words, which is expected and is in consistence with the results of other studies. However, the results are closer to Schmitt's estimates (2000) that there is not a large gap between receptive and productive knowledge of words, and do not confirm Eringa and Chamberlain estimates (as cited in Melka, 1997) that receptive knowledge may be several times higher than productive knowledge. The only significant difference between receptive and productive knowledge can be seen in the results of the fourth group, which may be due to the different type of lectures at the university level compared to learning English in elementary and secondary education.

Regarding the third question, the demonstration of word knowledge when the words are seen in isolation and in context, there are consistent higher results for words seen in context. It is obvious that context helps learners to recognize or infer the meaning of words. But the test for measuring productive knowledge in context demonstrates that when provided with context, learners can recall more words than when they are given in isolation. Therefore, we may conclude that the context triggers not only passive or receptive vocabulary, but productive vocabulary as well. What is more, the results of the fourth group indicate that older learners have even greater benefit from context, as they demonstrated the biggest difference in both receptive and productive knowledge when the words were seen in isolation or in context.

5. Conclusion

The aim of the present study was to determine the development of vocabulary knowledge in English language learners, specifically the vocabulary size at four stages of language learning. It also attempted to identify the difference between receptive and productive knowledge, and the difference in word knowledge when the words are seen in isolation and in context. We need to emphasize that the words included in the tests belonged to the first 5000 words, and words above that level and the words from the Academic Word List were not included. The results show that the learners constantly increase their vocabulary and that the greatest gain in vocabulary knowledge happens between the age of twelve and fifteen, that is after they have learned the most basic vocabulary during their first two years of learning English, their vocabulary knowledge rapidly increases during the next three years, and after that it increases at a slower pace. The results also show that the learners have greater receptive than productive knowledge of words, but the difference is not very big. And the last insight is that context is essential in helping learners recognize or recall words, especially for passive vocabulary, as the difference between receptive knowledge of words seen in isolation or in context is quite significant.

Webb (2005) believes that the difference between receptive and productive vocabulary is due to the fact that learners acquire vocabulary mostly through reading

and listening and because the classroom activities, such as matching words with their meanings or definitions, inferring words from context, or learning word pairs are more frequent than productive activities, like filling-in-gaps or writing tasks. As receptive exposure to vocabulary is not enough for gaining productive knowledge of words, Schmitt (2008, p. 345) suggests that learners should be equally engaged in productive activities in order to develop their productive vocabulary as well.

About the Author

Nina Daskalovska is associate professor at the Faculty of Philology, Goce Delcev University in Stip, Republic of North Macedonia, where she teaches courses on English language teaching methodology. Her areas of interest include ELT methodology, vocabulary acquisition, extensive reading, and computer assisted language learning. She has published more than a hundred journal articles, book chapters, textbooks and monographs.

References

- Folse, K. S. (2004). *Vocabulary Myths: Applying Second Language Research to Classroom Teaching*. Ann Arbor: University of Michigan Press.
- Goulden, R., Nation, P. & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, 11, 341-363.
- Harmer, J. (1991). *The Practice of English Language Teaching*. New Edition. New York: Longman.
- Hatch, E. & Brown, C. (1995). *Vocabulary, Semantics, and Language Education*. Cambridge: Cambridge University Press.
- Meara, P. (1996). The dimensions of lexical competence. In: G. Brown, K. Malmkjaer and J. Williams (Eds.) *Performance and Competence in Second Language Acquisition* (pp. 35-52). Cambridge: Cambridge University Press.
- Meara, P. (1997). Models of vocabulary acquisition. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 109-121). Cambridge: Cambridge University Press.
- Meara, P. (2002). The Rediscovery of Vocabulary. *Second Language Research*, 18(4), 393-407.
- Melka, F. (1997). Receptive vs. productive aspects of vocabulary. In Schmitt, N. and McCarthy, M. (Eds.). *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press
- Nation, I. S. P. (n.d.a). Vocabulary lists. Retrieved September 2020, from: <https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary-lists>
- Nation, I. S. P. (n.d.b). Measuring learners' present level of vocabulary knowledge (vocabulary resource booklet). Retrieved May 2010, from <http://www.vuw.ac.nz/lals/staff/paul-nation/nation.aspx>
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Rowley, MA: Newbury House.

- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Schmitt, N. (2008). Review Article: Instructed Second Language Vocabulary Learning. *Language Teaching Research*, 12(3), 329-363.
- Stoddard, G. D. (1929). An experiment in verbal learning. *Journal of Educational Psychology*, 20(6), 452-457.
- Waring, R. (1997). A comparison of the receptive and productive vocabulary sizes of some second language learners. *Immaculata (Notre Dame Seishin University, Okayama)*, 1, 53-68.
- Waring, R. (2002). Basic Principles and Practice in Vocabulary Instruction. *The Language Teacher*, 26 (07).
- Wilkins, D. (1972). *Linguistics in language teaching*. London: Arnold.

Creative Commons licensing terms

Authors will retain the copyright of their published articles agreeing that a Creative Commons Attribution 4.0 International License (CC BY 4.0) terms will be applied to their work. Under the terms of this license, no permission is required from the author(s) or publisher for members of the community to copy, distribute, transmit or adapt the article content, providing a proper, prominent and unambiguous attribution to the authors in a manner that makes clear that the materials are being reused under permission of a Creative Commons License. Views, opinions, and conclusions expressed in this research article are views, opinions and conclusions of the author(s). Open Access Publishing Group and European Journal of English Language Teaching shall not be responsible or answerable for any loss, damage or liability caused in relation to/arising out of conflict of interests, copyright violations and inappropriate or inaccurate use of any kind content related or integrated on the research work. All the published works are meeting the Open Access Publishing requirements and can be freely accessed, shared, modified, distributed and used in educational, commercial and non-commercial purposes under a [Creative Commons Attribution 4.0 International License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).