

Adult Learners' Acquisition of Vocabulary Knowledge from Explicit Listening: A Multidimensional Construct

Dr. Thamer Alharthi

Assistant Professor of Applied Linguistics, Department of European Languages and Literature, Faculty of Arts and Humanities, King Abdulaziz University, Jeddah, Saudi Arabia

Accepted 15 July 2016, Available online 18 July 2016, Vol.4 (July/Aug 2016 issue)

Abstract

To date there has been little empirical research exploring the effect of explicit, instructional learning by listening on the acquisition and retention of vocabulary, both receptively and productively. To address this issue, the current study observed the learning of two groups (one control and one experimental) of a total of 40 adult learners of English as a Foreign Language (EFL) in a tertiary classroom setting in respect of orthography, grammar and form-meaning knowledge, tested via receptive and productive measures. All participants were administered a pretest, an immediate posttest, and a delayed test over 4-week intervals. The results indicated that a large proportion of target words were learned both receptively and productively through instructional listening involving several aspects of vocabulary, i.e. orthography, grammar and form-meaning. In acquisition and retention tests the members of the experimental group demonstrated significantly more knowledge about the target words than the members of the control group. The impacts of each of these learning conditions on different lexical aspects are discussed in detail.

Keywords: Vocabulary, Listening, Instruction, Orthography, Grammar, Form-Meaning

1. Introduction

1.1 Statement of the Problem

The importance of vocabulary in language learning, teaching and proficiency is no longer in doubt. As Schmitt (2010) points out, the one thing that everyone involved in the learning process – students, teachers, material developers and researchers – agrees to is that learning vocabulary is an essential element of mastering a second language. The truth of this is commonly observed by seeing students carry around dictionaries and not grammar books. And many researchers cannot resist quoting Wilkins (1972, p. 111) that, "Without grammar very little can be conveyed, without vocabulary nothing can be conveyed." The literature provides much research supporting the primacy of vocabulary knowledge in second language competence.

It is widely accepted that incidental vocabulary learning, both via listening and reading, plays an essential role in vocabulary learning. Researchers have long been interested in the relationship between reading and vocabulary and have found vocabulary knowledge to be a significant determinant of reading success (e.g. Day, Omura & Hiramatsu, 1991; Horst, Cobb & Meara, 1998; Pitts, White & Krashen, 1989). However, very few studies have looked at the potential for vocabulary learning in the

context of L2 listening, Elley (1989) and van Zeeland and Schmitt (2013) being notable exceptions. The scarcity of research into vocabulary learning through listening is clearly expressed by Webb's (2016) statement that "research on incidental vocabulary learning through listening is an under-researched area in both the L1 and L2 contexts" (p. 135). A further complication for research is that the type of vocabulary knowledge needed for listening, speaking, reading and writing is not all the same (Nation, 2001). Several studies have determined the percentage of lexical items of spoken discourse a learner needs to know to achieve adequate comprehension (Adolphs & Schmitt, 2003; Bonk, 2000; Staehr, 2009; Hu & Nation, 2000). For instance, Schmitt (2008) notes that knowledge of 2000-3000 word families is required to understand everyday spoken discourse in English, with 95% coverage of this considered adequate. However, to achieve 98% coverage of everyday spoken discourse, knowledge of 6000-7000 word families is required. This massive amount of vocabulary knowledge cannot be learned when it is confined to two to three hours of EFL classroom instruction per week (Kang, 2015). It is likely that a similar philosophy of minimal EFL instruction via listening obtains elsewhere.

The status of English as an international language has not only led to it being taught widely at universities and schools in both English-speaking and non-English speaking

countries to Non-Native Speakers (NNS) of English, but has even led to English increasingly being the medium of instruction at selected university faculties in countries where English is not the native language of the local population, e.g. in Hong Kong and Singapore. The need to be able to understand spoken English, and ideally being able to engage in conversation with Native Speakers (NS) of English, is thus not only obvious but steadily increasing. The main difficulties in understanding spoken English experienced by the majority of learners of English as a Foreign Language (EFL) are these: the speed with which English is spoken by the NS of English, being confronted by unknown or a limited range of vocabulary, being unfamiliar with the topic of conversation (and thus the vocabulary employed), finding the accent employed (regional or social) by the native speaker difficult to understand if not incomprehensible, and hearing what is being said only once. All these factors share in common a lack of vocabulary knowledge and thus making this perhaps the most significant obstacle for EFL learners to engage in conversation (Chang & Read, 2006).

All these issues apply to Saudi Arabia, a country with a growing population, as a result of which class sizes in both school and university are always large, with 30-35 students the norm, thus usually much larger than EFL classes in the western world, and for this reason, genuine oral communication is not much used or encouraged in Saudi Arabia. Even when taking a university/college entrance exam, for example, by students seeking admission to the Department of European Languages and Literature (DELL) at King Abdulaziz University (KAU), the test comprises only reading comprehension, grammar rules and short examples of writing. Consequently, listening has never been a major component of the examination for entrance to the DELL. Thus, given the large number of EFL university students with limited exposure to English outside the context of formal study yet with an urgent need to improve their spoken language competence, research into listening as a source of language input is clearly important.

It is quite common that students encounter unfamiliar words in the text and materials for communication purposes as well as in the spoken input. Spoken discourse is often characterized by unclear articulation as well as word boundaries based on lexical insufficiency which means that spoken text hardly gives enough information for an EFL listener to guess the meaning of the word. This lack of clarity of spoken input makes word segmentation rather difficult task for the EFL listener. Emphasizing the challenge of the spoken text, Goh (2000) looked at listening problems of English and he found that failure to recognize known lexical items in the input is one of the most significant barriers to understanding mentioned by the listeners themselves.

Consequently, one might hypothesize that learners' knowledge of vocabulary will predict listening comprehension. However, knowledge of words by itself does not necessarily mean that learners will be able to

understand the linguistic information in the given text. Other factors also play a part in spoken language and negatively impact comprehension and hinder learning. Research has shown that factors such as text type (Shohamy & Inbar, 1991), first language listening competence (Vandergrift, 2006) and topic familiarity (Schmidt- Rinehart, 1994) affect listening comprehension. The interesting question then becomes how vocabulary is learned through listening? The primary concern of the present study is to explore precisely the vocabulary knowledge uptake from listening exposure.

1.2 Background to the study

Many L2 learners and their instructors view vocabulary as a crucial part of language learning. Therefore, it is important for language teachers to develop and apply appropriate learning tasks to achieve successful vocabulary learning by L2 learners. Consequently, researchers of language learning intend to explore why certain tasks are more effective than others in L2 vocabulary learning. While most of the research into vocabulary acquisition in academic contexts has been carried out in the field of reading (Bernhardt & Kamil, 1995; Horst, 2005; Laufer, 1992; Laufer & Ravenhorst-Kalovski, 2011; Nation 2001, 2006, 2014; Pigada & Schmitt, 2006; Qian, 1999, 2002; Ulijn & Strother, 1990; Waring & Takaki, 2003), there has as yet been insufficient research into L2 intentional vocabulary learning through listening (Elley, 1989; Brown, Waring & Donkaewbua, 2008; van Zeeland & Schmitt, 2013; Vidal, 2003).

There is little evidence that new lexical items can be acquired explicitly through exposure to spoken language. Elley (1989), for example, investigated 7-8 year-old children's L2 vocabulary learning from listening to a single story. Elley (1989) reported that children retained the meaning of 20% of target words after listening three times. The explicitness of instruction seemed to have been crucial for vocabulary acquisition to take place and that students with teacher explanation experienced gains of 40% knowledge of new words. It seems then, that the teacher explanation may contribute to much larger vocabulary learning gains than listening without teacher support for vocabulary learning. It is noteworthy, that Elley's (1989) study is a landmark of vocabulary learning via listening and one of few attempts to look at the earlier stages of first language vocabulary learning. However, Meara (1984) argues that vocabulary learning does not take place in similar ways, considering the range of learners' proficiency. Therefore, it seems necessary to explore further the effects of input, i.e. of spoken language on vocabulary learning with L2/EFL adult learners.

Vidal (2003) explored vocabulary acquisition from academic lecture presentations with 116 Spanish university students and showed significant vocabulary gains. In particular, he sought to investigate the types of word knowledge (e.g. academic words, technical words and low-frequency words) and their acquisition through

listening to academic lectures. The results of the one-month delayed posttest indicated that vocabulary knowledge increased at an average of 50% compared with the pretest measures. Students also made greater gains in technical terms than in academic and other low-frequency ones.

Using audio-assisted reading than listening alone, Brown, Waring and Donkaewbua (2008) examined incidental L2 vocabulary through reading while listening with 35 Japanese EFL learners at tertiary level. The researchers found that "the subjects were able to learn new words from context and that they learned most words in the reading-while-listening mode" (Brown, Waring & Donkaewbua, 2008, p. 156) followed by reading only and then listening only. In support of the results by Brown, Waring and Donkaewbua (2008), Webb and Chang (2012) implemented aural support during a reading approach in their study with 82 Taiwanese EFL learners and found that vocabulary learning gains are likely to be larger for audio-assisted reading condition than for reading alone.

An investigation of lexical coverage achieved via L2 listening by van Zeeland and Schmitt (2012) used frequency-based analyses to examine the percentage of vocabulary needed for learners to achieve a certain level of comprehension, and how much vocabulary was crucial for language success. In a recent study of experimental L2 learning, van Zeeland and Schmitt (2013) set out to explore incidental vocabulary acquisition through listening adopting a so-called dimensions framework, which comprises multiple aspects of vocabulary knowledge (form, grammar, meaning). The main purpose of their study was to compare incidental vocabulary growth of three vocabulary knowledge dimensions, namely form recognition, grammar recognition and meaning recall. Performance on multiple vocabulary tests showed that word form and grammar are learned relatively easily through L2 listening, while the acquisition rate of word meaning proved to be very small and the least effective. The authors concluded that in order for L2 listening comprehension to be lasting and a valuable method for vocabulary learning, learners may need to hear the target words as many as 15 times or more.

In sum, the results of the aforementioned studies reinforce the conclusion that L2 listening input can result in very considerable vocabulary gains, although the amount of learning appears to be less than might occur through reading or writing. This might be due to the fact that lexical coverage differs for the purpose of comprehension of written versus spoken input (Milton, 2009). Substantially, learners need a larger amount of vocabulary to reach the kind of coverage associated with good comprehension in written than in spoken text (Nation, 2006; Pellicer-Sánchez, 2015). To the best of the researcher's knowledge, apart from Elley (1989) and van Zeeland and Schmitt (2013), there seem to be no other studies on the explicit learning of vocabulary through listening. Hence, considering the importance of

vocabulary development particularly in EFL contexts and the noticeably sparse studies in this area, there is a need for further research to shed more light on the limitations of the previous studies and in particular to provide greater clarity of how vocabulary is learned through listening

1.3 The present study

The majority of the abovementioned studies either lacked ecological validity by limiting listening to an oral story text, they did not include a comparison group (Elley, 1989), or they had a rather small number of participants specifically in the delayed post-tests (van Zeeland & Schmitt, 2013) any of which would limit the generalisability of the findings. Furthermore, with the exception of van Zeeland and Schmitt's (2013) study, these studies merely intended to examine word meaning during listening, neglecting other aspects of vocabulary knowledge (Brown, Waring and Donkaewbua, 2008; Elley, 1989; Vidal, 2003). Since there is much more to vocabulary knowledge than knowing its meaning (Nation, 2001), exploring multiple facets of vocabulary knowledge through listening is likely to provide a more accurate picture of the extent of vocabulary learning. Although Elley's (1989) study is pioneering and informative, it measured only receptive vocabulary gains through a 20-item multiple-choice test with the target words supplied in both aural and written forms. However, incorporating productive vocabulary assessment could have provided further insight into the potential usefulness of listening for EFL learners' vocabulary use. The current researcher therefore believes that there is a need for more methodologically sound research on vocabulary learning through listening. As a first step toward this objective, the current study concentrated on two distinct multidimensional constructs: depth of vocabulary knowledge and receptive and productive vocabulary knowledge, and adopted a pre-test post-test experimental design to test the usefulness of vocabulary learning through listening. Brown, Waring and Donkaewbua (2008) argue that more uptake of vocabulary might occur if the listening treatment is conducted in shorter and more manageable sessions. Consequently, the current study sets out to empirically investigate how EFL learners acquire vocabulary through listening over a four-week time span. It examines multiple aspects of vocabulary knowledge, namely orthography, grammar and form-meaning by including receptive and productive vocabulary measures. The following research questions are addressed:

- 1) To what extent does listening contribute to EFL learners' vocabulary knowledge of orthography in immediate and delayed vocabulary recall tests?
- 2) To what extent does listening contribute to EFL learners' vocabulary knowledge of grammar and form-meaning in both immediate and delayed vocabulary tests of reception and recognition?

2. Method

2.1 Participants and setting

The 40 study participants, aged between 19 and 22 years, were enrolled in a 15-week, intensive academic English language (BA) program at KAU. The program involved content-based and task-based instruction, delivered over 17 hours per week. The participants comprised two complete EFL classes at a sophomore tertiary level. They had had at least six years of English study at junior and senior high schools. Their exposure to English was mostly confined to a set of curriculum, classroom and textbook materials. At the time of the study, they had been assigned to a listening and speaking course as part of their BA, whose objectives were to develop their listening skills and to assist them in increasing the size of their vocabulary to enable them to cope with the demands of subsequent more advanced courses. The participants were randomly assigned to one of two learning conditions: 20 participants were assigned to an experimental group who received their teacher's explanation of target words (as part of their vocabulary learning) and 20 participants were assigned to a control group who received regular classroom instruction, which did not have a component specifically devoted to vocabulary learning. To obtain an overall picture of the participants' vocabulary knowledge, a version of the Vocabulary Size Test (VST) (Nation & Beglar, 2007) comprising 1,000 and 2,000 word levels was administered. All participants mastered the 2,000 most frequent words in English and the two groups were comparable in terms of English competence since their scores on the VST did not differ significantly. Both the experimental and the control group had the same number of teaching hours per week (3 hours) and followed the same course content except for the teacher's explanation. It should be noted that the inclusion of a control group was intended to ensure that learning did not occur as a result of taking the vocabulary knowledge test and to gain an accurate measure of vocabulary learning from the treatment.

2.2 Study materials

The present study used a special textbook – *Pathways Listening, Speaking and Critical Thinking* (Chase & Johannsen, 2012) – which had been used to teach the study participants English vocabulary in their BA class. The learning material consisted of all the target items to be learned in their course; these were given in bold blue accompanied by their corresponding vocabulary exercises. The target words were thus embedded within the reading-listening texts, followed by contextualized exercises where students were required to fill in gaps in sentences, to supply the correct item when given a definition, and to match the target words with their

relevant definitions from a list of words. The disadvantage of not using authentic words is that there will always be some concern as to the ecological validity of the findings. Therefore, we cannot be certain whether the items would behave in the same context as authentic or real words. Seven reading-listening passages were selected from the first two units covering real-world content from one of the National Geographic series: *Kabuki; The Year of Living Dangerously; William Shakespeare; Boys and Girls Test Their Geography; Canadian Boys Win World Geography Contest; The King Penguin: Challenges to Reproduction; Orchid Question & Answer*. These texts were professionally developed from authentic sources and were recorded by NS of American English to sound as natural and clear as possible, and to expose the study participants to treatments that were ecologically valid. The analysis of the texts in terms of readability was done through software called RANGE on a website called "The Lexical Tutor". The purpose of employing this software was to ensure that the texts have reliability levels of 98% at the 2,000 word level, which according to Nation (2001) is the ideal level for students to gain adequate comprehension of a text to permit successful guessing of the meaning of vocabulary items. The total number of the running words in the texts assumed to be known by the participants was 1679 and the Flesch-Kincaid readability index was 98.9¹

In order to ascertain the target vocabulary items, all the seven texts and all vocabulary exercises related to these texts were examined. Authentic vocabulary items were used in the present study because some words (e.g. words that do not appear in bold) provided no learning value to the participants. In total, 40 target vocabulary items were included in the current study. The target items were selected according to their frequency of occurrence bands. For this purpose, the BNC-20 version of VocabProfile (Cobb's 2016) online software was used, which dissects the texts into frequency bands at each 1,000 word level up to and including the 20,000 word level. The distribution of the vocabulary items in the BNC-20,000 word bands is presented in Table 1.

Table 1: Lexical Profile at each BNC-20 Word Level

Freq. Level	Tokens (No.)	Cumulative coverage token (%)
K-1 Words	4 (10.00)	10.00
K-2 Words	14 (35.00)	45.00
K-3 Words	16 (40.00)	85.00
K-4 Words	5 (12.50)	97.50
K-5 Words	1 (2.50)	100.00

¹ The Flesch-Kincaid gives a value that ranges from 0 to 12 which represents a readability level from Kindergarten to Grade 12. Therefore, a score of 9 means that American students in 9th grade can read and understand the text.

Table 1 shows that the 40 target words were made up of 4 words from the 1,000 word level, 14 from the 2,000 word level, 16 from the 3,000 word level and 6 from other sources, i.e. words not on any of those lists. Therefore, 45% and 85% of the words in the texts were high frequency ones (that is, the word families from the 2,000 and 3,000 word levels were being introduced at a greater rate than from the other word levels). Since grammar was one of the aspects of vocabulary knowledge to be examined in the present study, it was necessary to decide what types of word classes should be selected. Based on the distribution of the target words in the target texts in terms of their grammatical category, a ratio of 8 nouns : 8 verbs : 4 adjectives was determined.

2.3 Perceived training in vocabulary learning

The control group was taught by a teacher obtaining an MA degree whereas the intervention session for the experimental group was carried out by the present researcher and was spread out over a 4-week period. Participants were directed to the target words in any given text and instructed to identify the words they are familiar with by giving a definition of the word either in their L1 or in English. For example, while listening to the texts in the first two chapters of the textbook, the researcher usually wrote the target words on the board, stopped the audio recording and then participants were trained to analyze the text, i.e. to pay attention to the key words and to guess their meanings. When the experimental group participants showed their full understanding of the listening texts, the researcher continued playing the recording straight through. When the researcher was able to confirm that all target items were clearly understood, the experimental group participants started doing vocabulary exercises. These exercises included discussing the dialogues and the listening texts, consulting words in dictionaries and doing role plays related to the dialogues. By contrast, the teacher in the control group did not intervene in the training process of the participants but only practiced and discussed simple concepts in English, thus teaching practices being observed by the researcher. In other words, the control group participants received the normal course instruction with no explicit instruction regarding vocabulary.

2.4 Dependant measurement instruments

The nature of the study is a quasi-experimental independent-groups design when comparing the successfulness of vocabulary learning through listening and retention at three aspects of vocabulary knowledge: orthography, grammatical and form-meaning knowledge between two groups of learners: experimental and control. Therefore, it involved three stages, pre-test, treatment, and post-test. In particular, tests were measured to both groups one week before the treatment

began, immediately one week after the experimental session and two weeks after the treatment. Figure 1 displays the design of the current study.

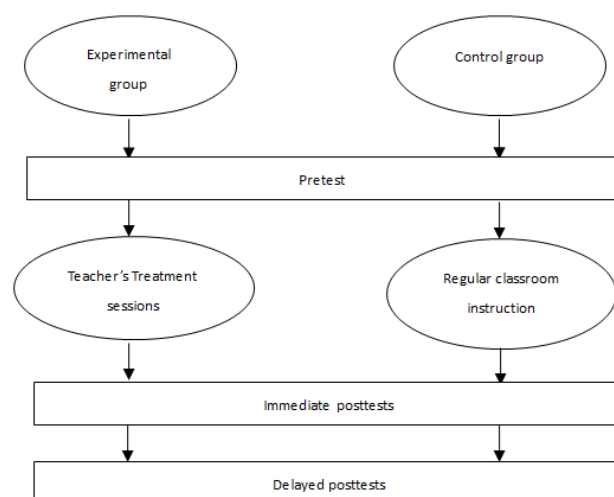


Figure 1: Study design

To scrutinize the depth of knowledge as well as receptive and productive vocabulary learning, the study employed three separate tests to measure aspects of vocabulary knowledge including orthography, grammar and form-meaning. The study adapted vocabulary measures from previous studies (Webb, 2009; van Zeeland & Schmitt, 2013). Most importantly the study used test formats for reasons of ecological validity as they were in line with participants' course materials. Therefore, participants were familiar with the test formats, i.e. gap-filling test, multiple-choice test and matching test.

To measure the orthography knowledge of vocabulary items, the researcher played a recording with 20 target words twice in the class. Participants had to supply the target words in a given sentence where the first letters of the target items were provided to ensure that the respondents would complete the correct word and not provide an alternative, for example:

They depend.....on their parents
 Instead of
 They defend on their parents.

This recall test format fits the purposes of the present study because it measures the kind of knowledge required for listening. That is, participants were given the phonological forms of the target words as a cue in their recall. Phonological cues were considered sufficient at least for the study participants since most of the rules of spelling were likely to have been learned by them.

A multiple-choice test was used to measure the grammar knowledge of vocabulary items. This test format was a standard, prompted receptive four-choice with the correct meaning and three distractors. All distractors were likely to be known by the participants as they consisted of words found in the participants' textbook

materials. The participants were given an *I don't know* option to reduce the effect of guessing. The researcher played the recording which included the target items and the participants were asked to choose the correct answer if they knew what part of speech the word was, i.e. noun, verb or adjective. Care was taken to ensure that the test contained short sentences with different parts of speech to allow the knowledge of the target word's part of speech to be demonstrated. After each item on the recording, participants were given ten seconds to read the sentences and choose one of the four options. If the participants heard *attract* on the recording, they would be able to choose (b) as the correct answer, as shown in the following example:

- a. The main **attraction** of the party was the good food.
- b. Many orchids have found fascinating ways to **attract** insects.
- c. It was an **attractive** painting.
- d. I don't know.

In the final test, the knowledge of form-meaning was measured using a recognition format, namely a matching test. This test also contained 20 target items with a space beside each item for a written response and each set of target words was divided into 10 blocks. Each block consisted of approximately 10 target words and one distractor and the participants were requested to choose the L2 meaning and / or definition of each target word. The inclusion of a distractor was to minimize the chances of successful guessing. An example of one block is illustrated below:

- | | |
|--------------------|--|
| 1) ____Possibly | a. a space or difference between two things. |
| 2) ____Gap | b. maybe. |
| 3) ____Knowledge | c. to arrive at an opinion or judgment. |
| 4) ____Investigate | d. to give someone confidence to do something. |
| 5) ____Equal | e. able to give your opinions strongly and clearly. |
| 6) ____Ability | f. a point in a certain scale, such as quality or ability. |
| 7) ____Level | g. to look at something carefully, research. |
| 8) ____Assertive | h. a skill that makes it possible to do something. |
| 9) ____Conclude | i. the things you know about. |
| 10) ____Encourage | j. the same in number, size or value. |
- k. something difficult that requires effort.

2.5 Procedure

The tests were carefully sequenced, hence the productive or recall task was administered before the receptive or

recognition tasks so the latter would not serve as cues to complete the productive test. Within a 4-week period and on different days, the participants were presented with a series of three dialogues and conversations. In the first phase, for the pretest, the participants in both groups completed the three vocabulary tests but no listening tasks were given. The purpose of the pretest was to assess the extent of word knowledge participants may have had prior to the treatment sessions. One day after the pretests, participants were assigned VST to control for any proficiency differences between the two groups as well as to distract the control group from focusing exclusively on the key words in the follow-up tests. The learning phase (treatment) was conducted within a week, comprising three periods of +/- 120 minutes, where the experimental group participants received formal teaching of the 40 items. While the experimental group participants took the same three tests immediately after the treatment, the control group participants completed the same three tests without being exposed to any treatment. The purpose of the immediate posttests was to determine whether explicit instruction of listening facilitated the participants' learning of new vocabulary knowledge. The delayed posttests were administered to both groups two weeks after the treatment sessions without any warning. It was essential to measure the participants in the same way in which they had learned so as to maintain reliability of data. Therefore, in each round of testing, the participants listened to the recording of the prompts on audio-CD and marked their answers on paper. The test items in each phase were kept the same but the key words were reordered to minimize transfer from the immediate to the delayed testing sessions. The input of all tests was presented in paper and pencil and all test sessions were unannounced.

2.6 Scoring and data analysis

All pretests and posttests were scored immediately using the same system. The test items were scored with one point for correct, and with zero points for incorrect, missing or *I don't know* answers. To examine vocabulary gains on the basis of listening between groups, relative learning gains were calculated for both groups adopting the following formula: [(immediate post-test score – pre-test score) / (total number of test items) × 100]. For the delayed post-test data, the same formula was applied: [(delayed post-test score – pre-test score) / (total number of test items) × 100]. The independent variables are treatment, dimensions of vocabulary knowledge (orthography, grammar and form-meaning) and the scores obtained from the tests are the dependent variables. In order to decide on the type of statistical analyses to be used for test score comparisons, One-Sample Kolmogorov-Smirnov Tests were conducted for the vocabulary scores of both control and experimental group. The results showed a normal distribution of data and therefore, the researcher decided to use parametric

analyses on both groups for pre-test, immediate test and delayed test. To determine the statistical significance of any learning that emerged between pre-test and immediate test, paired-samples *t*-tests were conducted. For each group, an ANOVA with repeated measures were performed to identify whether the gains from the pre-test and the tests conducted one week after treatment and two weeks after treatment respectively reached the level of significance (0.05). Each participant in this within-subject design reported three scores that were not totally independent of each other; repeated measures ANOVA

reducing error by taking this into account. All the inferential tests were done with the 0.05. alpha level of significance.

3. Results

Table 2 shows the descriptive statistics means, standard deviations, minimum and maximum scores, for the vocabulary tests (pre-test, immediate and delayed tests for both control and experimental groups). Figures 2-7 show the same results graphically.

Table 2 Descriptive statistics for vocabulary learning rates on the three tests

Condition	Type	Session	Mean	Minimum	Maximum	Standard deviation
Control	Orthography	Pretest	38.5	15	70	17.17
Experimental	Orthography	Pretest	33.7	10	55	14.58
Control	Orthography	Immediate	43.7	15	70	15.03
Experimental	Orthography	Immediate	66.7	45	85	13.4
Control	Orthography	Delayed	32	10	60	15.67
Experimental	Orthography	Delayed	67.7	40	80	11.03
Control	GM N	Pretest	46.8	12.5	87.5	24.62
Experimental	GM N	Pretest	48.7	12.5	87.5	22.17
Control	GM N	Immediate	53.1	25	87.5	19.06
Experimental	GMN	Immediate	82.5	50	100	17.86
Control	GM N	Delayed	36.8	12.5	75	17.05
Experimental	GM N	Delayed	84.3	37.5	100	18.08
Control	GM V	Pretest	43.1	12.5	87.5	21.64
Experimental	GM V	Pretest	43.7	12.5	75	19.23
Control	GM V	Immediate	55	25	100	19.61
Experimental	GM V	Immediate	72.5	37.5	100	21.3
Control	GM V	Delayed	33.7	12.5	75	17.25
Experimental	GM V	Delayed	81.2	37.25	100	18.36
Control	GM Adj	Pretest	63.7	25	100	25.06
Experimental	GM Adj	Pretest	38.7	25	75	17.15
Control	GM Adj	Immediate	73.7	25	100	22.17
Experimental	GM Adj	Immediate	83.7	25	75	23.33
Control	GM Adj	Delayed	47.5	25	75	19.7
Experimental	GM Adj	Delayed	86.2	50	100	19.07
Control	Form/meaning	Pretest	55	35	90	18.13
Experimental	Form/meaning	Pretest	46.5	25	65	11.82
Control	Form/meaning	Immediate	65	30	100	20.06
Experimental	Form/meaning	Immediate	82.2	55	100	12.4
Control	Form/meaning	Delayed	44.5	20	80	15.63
Experimental	Form/meaning	Delayed	83.2	50	100	14.26

Differences in the means between the control and the experimental group showed increases in vocabulary knowledge particularly from the pretest to the immediate test. However, the experimental group generally performed better on the immediate and delayed tests than the control group. When we closely examine the data in Table 2, it shows that the participants who had the most to learn based on treatment, made the most gains. Moreover, recognition gains were always higher than recall gains and this difference increased even one month after the instructional treatment ended. The experimental group participants scored highest on the grammatical receptive test followed by the form-meaning recognition test and the orthography recall test. In grammatical knowledge, they had a mean gain in respect of adjective of +47.5% while the averages for verbs and nouns were 37.5 and 35.6 respectively.

A paired sample *t*-test showed that the gain for the adjective category from pretest to immediate test was statistically significant ($df=19$, $t=-8.461$, $p < .001$). The mean relative gain in the form-meaning knowledge test was +36.7 %. The difference between the pretest and immediate test scores ($df= 19$, $t= -12.801$, $p < .001$) was relatively significant. As shown in Table 2, the relative gain from pretest to delayed test in orthography knowledge was +34%. A paired sample *t*-test confirmed the statistically significant difference ($df= 19$, $t= -9.266$, $p < .001$) between the pretest and immediate test scores. Consequently, the results suggest that treatments resulted in vocabulary learning in all aspects of knowledge and that vocabulary knowledge was retained in the long term.

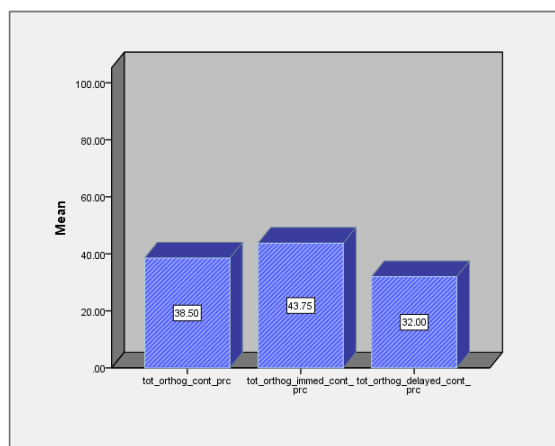


Figure 2 Orthography recall test for control group

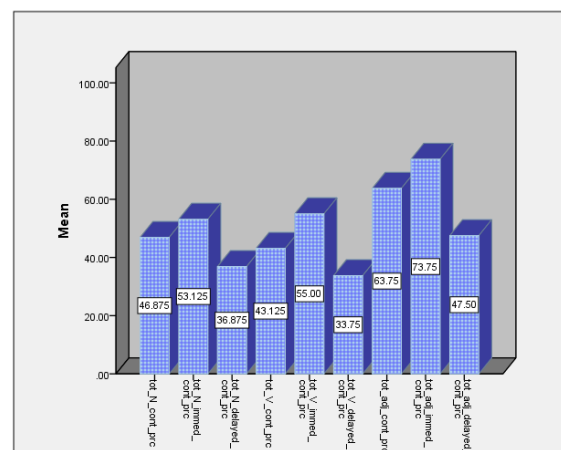


Figure 4 Grammatical receptive test for control group

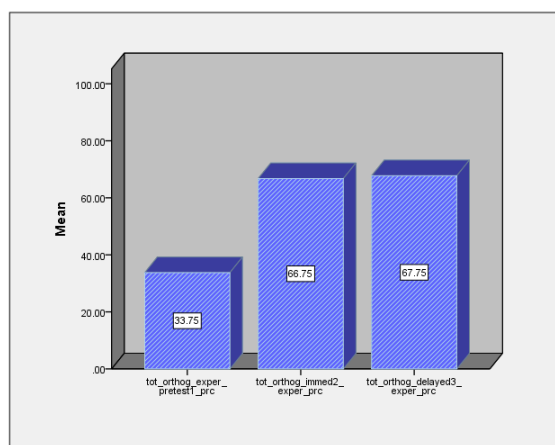


Figure 3 Orthography recall test for experimental group

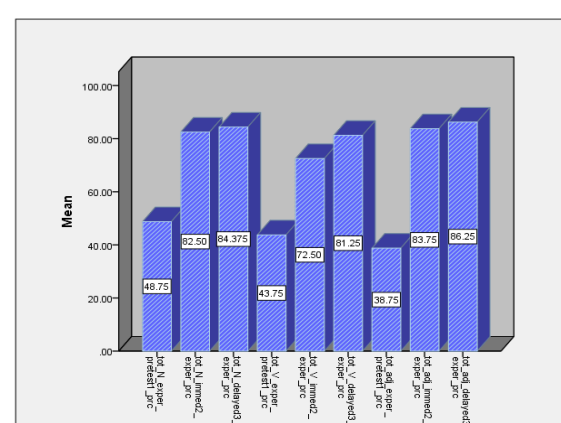


Figure 5 Grammatical receptive test for experimental group

Comparisons for all three dimensions of vocabulary knowledge showed no significant difference between the pretest means scores for the control and experimental groups, indicating that both groups had similar prior knowledge of the target words. However, the gap between the two groups became larger at the end of the study and the profiles displayed in Figures 2-7 describe this finding. The mean scores on the orthography recall test for the control group were ($M= 38.5$) and ($M= 43.7$) between pretest and immediate test, indicating a gain of +5.2% words. Despite the slight gain recorded, a paired sample t -test shows that the difference was statistically significant over the two test sessions ($df=19$, $t= -2.502$, $p < .022$). The mean scores on the grammatical receptive test were ($M= 46.8$) and ($M= 53.1$) for noun category; ($M= 43.1$) and ($M= 55.0$) for verb category; ($M= 63.7$) and ($M= 73.7$) for adjective category between pretest and immediate test, resulting in a gain of +6.3%, +11.9% and +10.0% respectively. However, paired sample t -tests did not reach the level of significance. The mean scores on the form-meaning recognition test were ($M= 55.0$) and ($M= 65.0$) leading to a gain of +10% between pretest and immediate test. The difference between the two test sessions was ($df=19$, $t= -2.571$, $p < .019$) notably significant.

Inspection of the estimated means displayed in Figures 2-7 indicated that there were considerable attrition over a 4-week period in the scores of all three tests which indicates that the experimental group participants again benefited from the treatment since their scores in the three vocabulary knowledge dimensions were twice as high as those of the control group participants. While a comparison of losses also seems to suggest that control group participants experienced a great loss of knowledge in the three dimensions of the delayed tests, the experimental group participants showed a significant degree of growth for all sessions. As can be gleaned from Figures 2-7, the scores obtained for both groups on the orthography recall test seemed to be far lower than would have been achieved simply by selecting the best answer on the grammatical receptive test as well as the form-meaning recognition test in all three test sessions. ANOVA tests were carried out to determine if there were any significant differences between the scores on the pretests and delayed tests for each knowledge type. A repeated-measures ANOVA revealed that the three dimensions of vocabulary knowledge were significantly different for the experimental group participants ($F= 11.644$, $p<.001$).

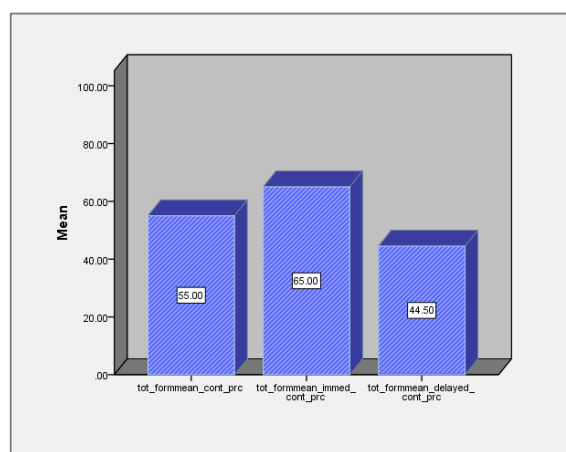


Figure 6 Form-meaning recognition test for control group

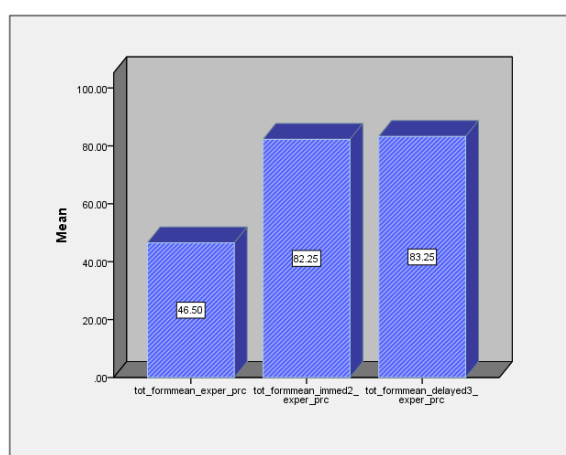


Figure 7 Form-meaning recognition test for experimental group

Although the repeated measures analysis showed that the effect of treatment was indeed significant ($F = 30.529$, $p < .001$) which was both linear and quadratic, the interaction between the three dimensions of vocabulary knowledge and treatment was not significant. The main effect of treatment supports the evidence that mean vocabulary gain scores by experimental group participants are not the same for all test sessions and that participants demonstrated much greater learning when they receive teacher's treatment sessions than normal classroom instruction.

4. Discussion

The present findings seem to indicate that although both groups exhibited gains in vocabulary knowledge through listening, the teacher's treatment sessions were a more efficient source of acquisition. This finding is consistent with Elley's (1989) study that found positive effects of instructional treatment on vocabulary acquisition and lent support to claims by Nation (2001) and Schmitt (2008) that explicit learning (i.e. attempting to learn the meaning of words intentionally) plays a much more prominent role

in L2 vocabulary acquisition. The experimental group participants showed significant improvement in all types of vocabulary knowledge while the control group did not. This was true for the immediate test, but more importantly it was also the case for the delayed test. Thus, the instructional treatments received by the experimental group were effective in facilitating durable learning. The data show that the participants learned some target words from the listening treatment in two types of vocabulary knowledge, but fewer words were learned in one type of word knowledge. This suggests that a considerable amount of vocabulary knowledge was retained on the grammatical receptive test and form-meaning recognition test but few words were retained on the unprompted orthography recall tests. This suggests that the recognition of words with teacher's explanation is learned before the meaning can be provided on the recall test. These findings seem to be compatible with results reported in studies that learners' receptive vocabulary tends to be larger than their productive vocabulary (Alharthi, 2014, 2015; Alharthi & Alfotais, 2015; Laufer & Paribakht, 1998; Webb, 2009). It also supports Nation's (2001) argument that productive or recall word learning is more difficult than receptive or recognition word learning.

The results revealed that experimental participants demonstrated orthographical knowledge of words immediately after the instructional treatment of +33% which was slightly lower than small-scale studies found that looked at vocabulary learning through listening or extensive reading. Vidal (2003) found a gain of 50.04%; van Zeeland and Schmitt (2013) found gains of 45.8% on receptive knowledge of form and Webb and Chang (2012) found a gain of 44.06% on the immediate posttest. This was the case for the easier recognition test formats employed in these studies. For instance, van Zeeland and Schmitt's (2013) study involved recognizing the form of the words through multiple-choice tests with target words presented in the list of options. Moreover, Webb and Chang (2012) used a bilingual matching test where their participants had to choose the L1 meaning for each target word. However, in the present study, the knowledge of orthography involved filling in gaps with missing letters, therefore participants had to reduce the number of possible word choices while listening to the sentence. The results of the delayed tests showed that the relative learning gains for orthography knowledge one month after the treatment were +34%. This gain is much higher than the learning gains found by van Zeeland and Schmitt (2013) whose durable learning from listening for form was 25% lower over a 4-week interval. The substantial increase between the preset and the delayed test for orthography knowledge found in the present study might be accounted for by the explicit vocabulary learning, i.e. the instructional treatment, which could indeed be an appropriate and effective way to bring about such knowledge gain in a classroom-based course. The current study adds to the growing body of evidence

supporting this argument. One might speculate that the participants in the present study might have looked up the target words they saw on subsequent tests, or they might have encountered some of the words in subsequent input (e.g. at home or in other classes). It turns out that the increase is attributable to external factors and may not be regarded as retention based on instructional treatment. Such an explanation is inadequate, however, since one can never rule out that participants after submitting their immediate tests checked and verified their answers with their fellow students.

The similarity between the types of measures used in the present study may result in little difference between the mean scores of different aspects of vocabulary knowledge. The size of the gains in grammatical and form-meaning knowledge indicates that the different aspects of knowledge may develop together over the course of study. In particular, most of the learning took place in grammar which was 47.5% in the adjective category over a 4-week period. These results contrast with those by van Zeeland and Schmitt (2013) who found that between the pretest and delayed posttest there was an increase in the recognition of grammar of 24.6% and for recall meaning of 7.5%. The fact is that instructional treatment almost always holds an advantage in vocabulary learning. These findings are in line with the conclusion reached by Spada and Tomita (2010) that treatment leads to higher grammatical knowledge gains than the lack of it. More encouragingly, the data from the type of instruments used to assess vocabulary gains, namely multiple-choice test and matching test, had a great bearing on retention rates. Therefore, partially known words that were not measured through the recall test (i.e. filling in gap) were likely known by the prompted receptive and recognition tests (i.e. multiple-choice and matching tests). One should emphasize that one element which likely led to this good learning is the effect of treatment which may well have helped the experimental group participants to consolidate the learning achieved in the listening and explicit exercises. Such environments would likely result in the negotiation of unknown knowledge of a form-meaning link which in turn facilitates learning (Newton, 1995; Webb, 2005; Webb & Change, 2012). Yet, it suggests that repeated tests at subsequent intervals can act as useful recycling devices in the vocabulary learning process. Taken as a whole the current study indicates that both receptive and productive learning can be very effective methods of gaining multiple aspects of vocabulary knowledge through listening treatment. The findings suggest that if the purpose of learning is to develop comprehension, receptive and recognition tests may be more effective, and if its purpose is to improve production, recall may be better (Webb, 2009).

Conclusion

The present study shows that a considerable amount of vocabulary knowledge was gained from classroom

listening instruction. The percentage of vocabulary learning through listening instruction appears to be much higher for both receptive grammatical knowledge and recognition of form-meaning than for recall of orthography knowledge. It appears therefore that some types of vocabulary knowledge, e.g. grammar and form-meaning, are learned relatively easily with the support of the teacher's explanation, while others, e.g. orthography, are not. This intervention study has shown that relatively few new lexical items are learned from listening as measured by filling in a gap test. On the other hand, more vocabulary knowledge is learned and retained from listening if multiple-choice and matching tests are used as measures of vocabulary knowledge. As promising as these results are, the limited number of target items per condition requires a degree of caution in interpreting and generalizing from the results. However, it was clear that the type of treatment investigated permits participants to establish different aspects of word knowledge at receptive and productive levels. It would be useful for future research to look at the effect of listening instruction on vocabulary learning with aspects related to language use such as registration and formulaic sequences. Although the two cohorts in the current study had similar EFL learning profiles, e.g. age, lexical proficiency, taking the same course at the same grade, female students would likely react to the instructional treatments differently, and perform differently, from their male counterparts which is an issue deserving of further attention.

References

- [1]. Adolphs, S. and Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics* 24, 4, pp: 425–438.
- [2]. Alharthi, T. (2014). The impact of attrition on vocabulary knowledge among Saudi teachers. *English Language Teaching*, 7(4), pp: 111–118.
- [3]. Alharthi, T. (2015). Adding More Fuel to the Fire: A Study of Attrition in Formulaic Sequences by Adult Learners. *Arab World English Journal*, 6(3), pp: 230–243.
- [4]. Alharthi, T, & Alfotais, A. (2015). Textbook Input as Agent of Low Vocabulary Uptake by EFL Learners. *Life Science Journal*, 12(10), pp: 1–7.
- [5]. Bernhardt, E, & Kamil, M. (1995). Interpreting Relationships between L1 and L2 Reading: Consolidating the Linguistic Threshold and the Linguistic Interdependence Hypotheses. *Applied Linguistics*, 16,1, pp: 15–34.
- [6]. Bonk, W. J. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening*, 14, pp: 14–31.
- [7]. Brown, R., R. Waring & S. Donkaewbua (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language* 20.2, pp: 136–163.
- [8]. Cobb T (2016) Web VocabProfile. Available at: www.lexutor.ca
- [9]. Chang, C-S., Read, J., 2006. The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly* 40, pp: 375–397.

- [10]. Chase, B.T and Johannsen, K.L. (2012). *Pathways 3 Listening, Speaking and Critical Thinking*. Boston MA: National Geographic Learning.
- [11]. Day, R., Omura, C., & Hiramatsu, M. (1991). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language*, 7, pp: 541–551.
- [12]. Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, 24, pp: 174–187.
- [13]. Goh, C. (2000). A cognitive perspective on language learners' listening comprehension problems. *System* 28, pp: 55–75.
- [14]. Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *The Canadian Modern Language Review*, 61, pp: 355–382.
- [15]. Horst, M., Cobb, T., & Meara, P. (1998). Beyond a Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11, pp: 207–223.
- [16]. Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), pp: 403–430.
- [17]. Kang, E.Y. (2015). Promoting L2 Vocabulary Learning through Narrow Reading. *RELC*, 46(2), pp: 165–179.
- [18]. Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126–132). London: Macmillan.
- [19]. Laufer, B., & Paribakht, T.S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48, pp: 365–391.
- [20]. Laufer B, Rozovski-Roitblat B (2011) Incidental vocabulary acquisition: the effects of task type, word occurrence and their combination. *Language Teaching Research* 15(4): pp: 391–411.
- [21]. Meara, P. (1984). 'The study of lexis in inter-language' in A. Davis, C. Cripe, and A.P.R. Howatt (eds): *Interlanguage*. Edinburgh: Edinburgh University Press. (pp.225–35).
- [22]. Milton, J. 2009. *Measuring Second Language Vocabulary Acquisition*. Bristol, UK: Multilingual Matters.
- [23]. Nation, I. S. P. (2001). *Vocabulary learning in another language*. Cambridge: Cambridge University Press.
- [24]. Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, pp: 63, 59–82.
- [25]. Nation, I.S.P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language*, 26, pp: 1–16.
- [26]. Nation, I.S.P. & D. Beglar, (2007). A vocabulary size test. *The Language Teacher*, 31, pp: 9–13.
- [27]. Newton, J. (1995). Task-based interaction and incidental vocabulary learning: A case study. *Second Language Research* 11, pp: 159–177.
- [28]. Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18(1), pp: 1–28.
- [29]. Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language*, 5, pp: 271–275.
- [30]. Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, 56, pp: 282–307.
- [31]. Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52, pp: 513–536.
- [32]. Staehr, L. (2009). Vocabulary knowledge and advanced listening comprehension in English as a Foreign Language. *Studies in Second Language Acquisition* 31, pp: 1–31.
- [33]. Pellicer-Sánchez, A. Learning L2 collocations incidentally from reading. *Language Teaching Research*, 3, pp: 1–22.
- [34]. Schmidt-Rinehart, B. (1994). The effect of topic familiarity on second language listening comprehension. *Modern Language Journal*, 78, pp: 179–189.
- [35]. Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12, pp: 329–363.
- [36]. Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke, England: Palgrave Macmillan.
- [37]. Shohamy, E., & Inbar, O. (1991). Validation of listening comprehension tests: The effects of text and question type. *Language Testing*, 8, pp: 23–40.
- [38]. Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60, pp: 263–308.
- [39]. Ulijn, J. M., & Strother J. B., (1990). The effect of syntactic simplification on reading EST texts as L1 and L2. *Journal of Research in Reading*, 13, pp: 38–54.
- [40]. Vandergrift, L. (2006). Second language listening: Listening ability or language proficiency? *Modern Language Journal*, 90, pp: 6–18.
- [41]. van Zeeland, H. & N. Schmitt (2012). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics* 34.4, pp: 457–479.
- [42]. van Zeeland, H. & N. Schmitt (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System* 41.3, pp: 609–624.
- [43]. Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, 24(1), pp: 56–86.
- [44]. Waring, R. and Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15, pp: 130–163.
- [45]. Webb, S., (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, 27, pp: 33–52.
- [46]. Webb, S. (2009). The effects of pre-learning vocabulary on reading comprehension and writing. *Canadian Modern Language Review*, 65, pp: 441–470.
- [47]. Webb, S., & Chang, C.-S. (2012). Vocabulary learning through assisted and unassisted repeated reading. *The Canadian Modern Language Review*, 68, pp: 276–290.
- [48]. Webb, S. (2016). Replication Studies: Learning vocabulary through meaning-focused input: Replication of Elley (1989) and Liu & Nation (1985). *Language Teaching*, 49(1), pp: 129–140.
- [49]. Wilkins, D. A. (1972). *Linguistics in language teaching*. Arnold: London.