



Journal of Teaching Language Skills (JTLS)
38(4), Winter 2020, pp. 165-203- ISSN: 2008-8191
DOI : 10.22099/jtls.2020.37154.2821

The Effect of “Narrow Reading” on Learning Mid-Frequency Vocabulary: The Role of Genre and Author

Elaheh Sotoudehnama*

Marie Virginia Raye Ahmadi **

Maryam Asadi Zarmehri ***

Abstract

This study investigated the effect of Narrow Reading (NR) on learning mid-frequency words. Vocabulary Size Test (VST) designed by Nation and Beglar (2007) was administered as the first pre-test to 196 students, from among whom 91 students whose vocabulary size ranged between 2100-3500-word families, $\pm 1 SD$, became the target of this study and were randomly classified into four groups. Then another pretest administered was a modified version of Wesche and Paribakht’s Vocabulary Knowledge Scale (VKS) which assessed how well they knew twenty vocabularies derived from the expository and narrative texts. After thirty sessions for direct instruction, VKS was repeated immediately as a posttest to assess vocabulary learning while a month later, VKS was taken as a delayed posttest to assess vocabulary retention. To address research questions, a mixed design factorial ANOVA was conducted. The findings revealed that while the theme was fixed, reading expository texts rather than narrative ones led to the growth in vocabulary knowledge. In addition, the group that read narratives forgot most of the mid-frequency words in the retention phase. As opposed to multi-author texts, reading texts written by one author resulted in a better vocabulary achievement in both learning and retention phases. It was also revealed that there was no interaction

Received: 04/05/2020

Accepted: 31/08/2020

* Professor, Department of English Language and Literature, Faculty of Literature, Alzahra University - Email: esotoude@alzahra.ac.ir, Corresponding author

** Assistant Professor, Department of English Language and Literature, Faculty of Literature, Alzahra University - Email: mvahmadi@alzahra.ac.ir

*** Graduate student, Department of English Language and Literature, Faculty of Literature, Alzahra University - Email: maryamasadi262@gmail.com

between genre and author. In other words, the author was as influential in learning mid-frequency words among expository texts as it was in narrative ones. Finally, it was concluded that NR has a positive effect on learning mid-frequency words.

Keywords: Narrow reading, Theme, Vocabulary

During the past 30 years, the act of reading has been investigated for various reasons. Among these investigations, the development of vocabulary through reading has been confirmed (Elley, 1991; Jenkins, Stein, & Vygotski, 1984; Krashen, 1989; Nagy, Anderson & Herman, 1987; Saragi, Nation & Meister, 1978; Stanovich, 2000). Another reason for this research was that reading is not only a means of gaining knowledge but also an entertainment source. Thus, to achieve these two objectives, different types of reading have been utilized. In the past, approaches such as intensive reading or extensive reading (ER) were considered effective in the context of a second or foreign language, but these have been challenged. For instance, Alderson and Urquhart (1984) claimed that intensive reading is just a language exercise. It is a tool to practice language aspects such as syntax and semantics rather than reading itself. On the other hand, extensive reading, which is defined as reading large amounts of material in the second language, exposes readers to a wide range of vocabularies, and hence the opportunity of exposure to familiar words is infrequent (Nation, 1997).

To compensate for these limitations, Krashen (1981) introduced “Narrow Reading”. It is said that books written by the same author, considering a common genre or theme, would ease reading since reading such books provide readers with familiar context (Cho, Ahn, & Krashen, 2005). In other words, the more one reads in one area, the more s/he would be informed with that area, and comprehends the text easier. What is more, narrow reading (NR) exposes learners repeatedly to vocabulary and grammar, which results in stabilizing the vocabularies in students’ minds (Krashen, 1981). Also, Krashen (2004) asserted that “since each writer has favorite expressions and

a distinctive style, and each topic has its own vocabulary and discourse, narrow reading provides built-in review” (p. 17). Although these texts are not controlled for vocabularies, reading a series of books written by one author or on the same topic increases the repetition of words; thus, understanding the texts will be facilitated. Nevertheless, not only are there a few studies to confirm the effectiveness of it but also each of them emphasizes only on one of the elements introduced by Krashen (Ahmadi & Poordaryaie, 2015; Cho, Ahn & Krashen, 2005; Gardner, 2008; Hansen & Collins, 2015; Kyongho & Nation, 1989; Schmitt & Carter, 2000).

Word Frequency

One of the most important factors which affect identification, recognition, and recall of a new word is word frequency (Gorman, 1961; Hall, 1954; Sumbly, 1963). Traditionally, vocabularies were classified under two categories, i.e., high and low-frequency vocabularies. The first 2000-word families were referred to as high frequent while over 2000-word families up to all of the word families beyond the 10000-frequency level were called less frequent. Later, Cobb (2007) challenged such categorization by claiming that the first 3000-word families in Nation’s BNC list represent the most frequent words.

In 2014, Schmitt and Schmitt proposed that low- frequency word families should be moved from 10000 level to the 9000 level since these 9000-word families lead to the understanding of at least 95.5% of the texts. To highlight the ratio of high-frequency words to low-frequency, Nation (2001) conducted an analysis. His analysis of the vocabularies of the American Heritage Intermediate revealed that the first 1000-word families cover 74.1% of the running words. Conversely, less frequent words cover 25.9% of running words. Therefore, there is no doubt that the number of words that occur very frequently in the texts is small. In contrast, there is a large number of words that occur less frequently (Schmitt & Schmitt, 2014).

Mid-frequency Words

Schmitt and Schmitt (2014) claimed that “high-frequency vocabulary in English extends up to 3000-word families and, that low- frequency ones begin at about the 9000-frequency level” (p.494). This left a great gap between the 3000 and 9000 level which was labeled as mid-frequency words. The fact was that there was a dearth of research examining the importance of learning mid-frequency words (Folse, 2010; Horst, Collins & Cardoso, 2009; Nation, 2013; Schmitt & Schmitt, 2014; Webb & Rodger, 2009a, 2009b).

Laufer (1996) claimed that knowledge of 3000-word families leads to understanding 90% to 95% of texts. However, according to Webb and Rodgers (2009a, b), knowing 3000-word families results in understanding about 95.45% of movies and television programs. On the other hand, Nation (2013) announced that learning the mid-frequency words would increase the reader’s comprehension of a text up to 98%. Schmitt and Schmitt (2014) mentioned that since mid-frequency words are not addressed in class, they would be easily forgotten. High-frequency words are taught but mid-frequency words are assumed to be picked up from exposure (Horst, Collins, & Cardoso, 2009). Thus, lack of frequent repetition and direct teaching turns out to be problematic in learning mid-frequency words (Folse, 2010).

The Impact of Word Encounter on Retention

The scarcity of low-frequency vocabularies in the discourse has led scholars to conduct different researches to enable learners to absorb them. One of them which drew attention was the impact of exposures on vocabulary learning. Some of the researchers claimed that the more readers encounter a word, the better they will remember it (Horst, Cobb & Meara, 1998; Jenkins, Stein & Wysocki, 1984; Nation, 1982; Perry & Mc Donald, 2001; Pigada & Schmitt, 2006; Saragi, Nation & Meister, 1978; Waring & Takaki, 2003). On the contrary, others believed that too many encounters would prevent learning since students no longer pay attention to those words (Webb, 2005).

The problem was that there was no consensus over the number of encounters. In 2010, Schmitt asserted that even one exposure results in measurable learning of not only word form but also its meaning. Still, he believed a few exposures are just sufficient for word recognition rather than its recall. He claimed that learning begins when students encounter a word 5 to 8 times and increases with 10 to 17 encounters. Although it is not clear how many exposures are needed to learn vocabularies, the point that meeting words happens through reading texts should not be ignored. However, the question is which reading comprehension approach is more effective in achieving such a goal.

Narrow Reading

The term NR was first coined by Krashen in 1981. Cho, Ahn, and Krashen (2005) defined NR as “reading in only one genre, one subject matter or the work of one author” (p.58). Considering this definition, one might say that one of the justifications for applying NR is that each writer has a set of admirable phrases that form his style, and each topic has some related expressions and words. Thus, NR leads to encountering these vocabularies repeatedly (Krashen, 2004). In fact, as a result of high word reiteration, implicit learning of vocabularies can be facilitated (Hunt & Beglar 2005; Kang, 2015; Schmitt & Carter, 2000). Next, when students read related texts, they will become familiar with the context. The fact is that the more they read narrowly, the better they comprehended the text (Yang, 2001). Third, narrow reading leads to recycling not only frequent words but also those words that are less frequent (Kyongho & Nation, 1989). Finally, the most distinguishing feature of narrow reading is that it does not intentionally control texts for the occurrence of words (Gardner, 2008). In fact, by applying authentic texts written by the same author, or having the same theme, the chance of vocabulary occurrence from different frequencies would be provided.

Although the concept of NR was proposed over 36 years ago, there were few studies conducted in this area. In one study, Gardner (2008) theorized that the effect of NR on the retention and recycling of vocabularies depended on the genre. After scanning the texts and listing words that were not high-frequency, he theorized that the number of authors made a difference in occurrence and recycling vocabularies of narrative texts while the presence of a specific theme seemed to be important in recycling vocabularies of expository texts.

Unlike Gardner's (2008) study on NR, which was theoretical, some others were conducted empirically. Ahmadi and Poordaryaie (2015) performed NR among three groups with different tasks of reading from a single author to multiple authors on a specific topic to various topics from different authors. The results of this study demonstrated that reading books by the same author was more efficient in terms of vocabulary breadth improvement. The problem with this study was that the chosen books were not authentic i.e., their vocabularies were manipulated to be comprehensible.

Schmitt and Carter (2000) compared the vocabularies of two types of texts: related compared to unrelated. Their survey indicated that the number of new vocabularies which occurred in related stories was less than unrelated ones. They concluded that "narrow reading facilitates reading by lowering the lexical load required of the learner" (p. 8).

Kang (2015) investigated the effects of NR on vocabulary expansion. Two classes who knew 1800 to 2000-word families participated in this study. The first group was labeled as a narrow reading group and the second as a regular reading group; both groups started to read a brief excerpt from a book. However, the narrow reading group continued to read related articles while the other one read passages on different subjects. The results indicated that the narrow reading group performed better than the control group in both receptive and productive vocabulary measures.

Paya-Guerrero and Segura (2015) compared the acquisition of vocabulary through NR in contrast to reading plus vocabulary enhancement activities. Although both groups improved in vocabulary learning, the narrow reading group outperformed the other one considering receptive knowledge. Since the two groups did not have the same vocabulary level in the pretest, one may conclude that this condition might have influenced the final results.

Abdollahi and Farvardin (2016) studied the effect of NR on EFL learners' vocabulary recall and retention. Although the results of this study revealed that the narrow reading group outperformed the control group in receptive vocabulary knowledge, one may suspect that an extraneous variable might have influenced the results since the target words were selected from 504 Absolutely Essential Words rather than the texts.

In conclusion, the existing researches in the area of NR were either theoretical or focused on one of NR elements such as theme, authorship, or genre. In this study, there has been an attempt to fill the gap of shortcomings as much as possible.

The Present Study

The studies in the area of NR were rare, but those that were conducted confirmed the importance of one of the elements (theme, author, or genre) raised by Krashen (1981). As an example, some of the investigations emphasized theme-based narrow reading, i.e., thematically related texts were used in their studies (Abdollahi & Farvardin, 2016; Kang, 2015; Kyongho & Nation, 1989; Paya-Guerrero & Segura, 2015; Rodgers & Webb, 2011; Satarsyah, Nation & Kennedy, 1994; Schmitt & Carter, 2000). On the other hand, there was only one study which stated that author-limited texts were the most influential way to expand vocabulary breadth (Ahmadi & Poordaryaie, 2015). On the contrary, Gardner's (2008) corpus-based research focused on genre as the determining factor. In this case, narrative texts written by one author led to the repetition of names (places, characters), while expository

texts having the same theme were more efficient in recycling thematic vocabularies that carry the content of the text. The researchers' purpose is to investigate the role of NR while considering all the elements mentioned by Krashen. In other words, in this study, the role of genre and authorship with the specified theme on increasing vocabulary knowledge was investigated. Thus, the study intended to answer the following questions:

1. Does genre (narrative compared to expository) play any significant role in vocabulary learning written by one author compared to multiple authors?
2. Does the number of authors (one author compared to multiple authors) play any significant role in vocabulary learning in narrative compared to expository texts?
3. Is there any significant difference between pretest, posttest, and delayed posttest scores among learners who learn mid-frequency words through reading expository compared to narrative texts written by one author compared to multiple authors (as reflected in the depth of learned vocabulary)?

The related null hypotheses to the research questions were formulated.

1. Genre (narrative compared to expository) does not play any significant role in vocabulary learning written by one author compared to multiple authors.
2. The number of authors (one author compared to multiple authors) does not play any significant role in vocabulary learning in narrative compared to expository texts.
3. There is not any significant difference between pretest, posttest, and delayed posttest scores among learners who learn mid-frequency words through reading expository compared to narrative texts written by one author compared to multiple authors (as reflected in the depth of learned vocabulary).

Method

Participants

About 164 non-English major students volunteered to take part in English classes for free at the dormitory of a state University in Tehran during winter and spring of 2017. The main criterion for the classification of students was their vocabulary size. Although students with different vocabulary sizes took part in this study, only students whose vocabulary size ranged between 2100-3500 were targeted, i.e., 91 students. These students were selected because their standard scores comprised the normal part of the curve, i.e., one SD above or below the mean. In other words, they knew the high-frequent word families (below 3000) and were ready to learn the mid- frequent ones.

Instruments

Three instruments were used in this study. These instruments include a sample of lexical profile (2015), vocabulary size test (2007), and Wesche and Paribakht’s measurement scale (1996).

Lexical Profile

Lexical profile updated in 2015 is a program developed on the basis of British National Corpus. It is available at the researchers’ disposal on Tom Cobb’s site (<http://lextutor.ca>). This program calculates the percentage of word tokens, families, and types in a text. Each level consists of 1000 word families. Words that are not in the frequency list appear in the “off-list” words.

In this study, the researchers used the lexical profile to classify words into different frequency levels. Moreover, it was assumed that the target group was ready to learn this range since their vocabulary size results indicated that they knew most of the high-frequency words. Therefore, it aimed to investigate whether narrow reading could lead to the expansion of vocabulary size or not. Thus, the lexical profile was applied to choose a list of 20 words categorized

as 3000-word families (Appendix A, B, C, & D). These words were incorporated in the test since they were repeated at least five times in texts.

Vocabulary Size Test (2007)

From among the three vocabulary size estimates which use either word family or lemma as the counting unit- X_Lex test developed by Meara and Milton (2003), VLT designed by Schmitt, Schmitt, and Clapham (2001), and VST suggested by Nation and Beglar (2007), the last one was used in this study due to the shortcomings of the first two and VST was used as a proficiency measure which determined how much vocabulary students knew. Vocabulary Size Test (VST) was used to measure the breadth of vocabulary knowledge. It is a standardized test that is highly reliable as indicated by Rasch reliability indices > 0.96 (Beglar, 2010). This test consists of fourteen levels, which begin with 1000 words and end with 14000 words. Each level consists of 10 items and 40 choices. Test takers were given the word and asked to match it with the most appropriate choice. According to Nation (2012),

A learner's total score on the 140-item test needs to be multiplied by 100 to find the learner's total vocabulary size. Thus, a score of 30 out of 140 means that the learner's vocabulary size is 3000-word families. (p. 12)

Based on the personal correspondence, there was no research which had correlated VST scores with CEFR levels or any scale describing language ability (Beglar, personal communication, June 26, 2017). In contrast to other vocabulary size estimates, VST could be used as a proficiency measure to assess receptive knowledge. Consequently, it was used as a pretest to determine how many words family students knew. The focus of this study was on the normal part of the curve, whose vocabulary size was about 2100-3500 based on the number of words they knew. This means that they knew most of the high-frequency words. Therefore, they were ready to learn mid frequent words.

A Modified Version of Wesche and Paribakht Measurement Scale (1996)

Wesche and Paribakht's VKS (1996) asserted that this measurement scale would assess the depth of vocabulary knowledge i.e., how well students know the meaning of vocabularies and their usage in context. Wesche and Paribakht reported high correlations (0.92 to 0.97) between the learners' responses and the way their responses were scored (cited in Waring, 2002). In this study, Wesche and Paribakht's (1996) was used as pretest, posttest, and delayed posttest. As a pretest, it enabled the researchers to diagnose those mid frequent words which are used in the text but unfamiliar to the students. As a posttest, the same test was administered to see if the exposure to mid-frequency had led to learning words, particularly, the 3000-word families. After a month, students took this test as a delayed posttest to examine how many mid-frequency words they had retained. A month interval was selected as Mackey and Gass (2005, p.149) suggested, “Often this is one week following the first posttest and then two weeks later and even two or three months later”.

Design

According to Hatch and Farhady (1981), “the concept of experimental design is an idealized abstraction. The reason is that we are dealing with most complicated of human behaviors” (p. 23). Since students took the tests voluntarily, it was not a random selection. In addition, only students who knew 2100-3500 word families were selected as the target group among those registered to participate in classes. Thus, a quasi-experimental design was used.

Data Collection Procedures

In the beginning, the researchers issued an advertisement requesting volunteers in the dormitory of Alzahra University. Since the main criterion for the selection of the target group was vocabulary size, the first pretest used in this study was a sample of vocabulary size tests (VST) that was devised by

Nation and Beglar (2007). It was administered to assess students' breadth of vocabulary knowledge.

To select the target group by VST, a series of actions were taken. At first, VST was administered to 164 volunteered students. Then, according to Nation (2012), students' scores were multiplied by 100 to determine how many word-families they knew. Next, students' scores were ranged from the highest to the lowest to calculate their standard scores. By applying the Z-score formula, it became obvious that 68% of scores fell between $\pm 1 SD$ which meant that they knew 2100-3500 word families. As a result, the focus of this study was on this group because they formed the normal part of the curve. Still, other students were not excluded from the class.

Group A studied two narrative books written by one author while two books by two different authors in the narrative genre by the group B. Then again two books written by one author in the expository genre were studied by the group C and two books by two authors in the expository genre by the group D. As an example, "The Curse of the Mummy's Tomb" and "Return of the Mummy" written by R.L. Stine was allocated to group A, while "The Great Egyptian Grave Robbery" a book written by Jeff Brown, and "The Shining Pyramid" written by Arthur Machen were used for group B. The other two groups read topic related expository books written by one author compared to multi-authors. For instance, "Pyramids" and "The Mystery of Egyptian Mummy" written by Joyce Filer were read by group C, whereas "The Tomb of King Tutankhamun" edited by Michael Woods and Mary B. Woods, and "Mummies and the Pyramids" by Sam Taplin was defined for group D. However, there were some reasons behind choosing these books. First and foremost, the main purpose of this study was to keep the theme fixed to see the effect of NR through different genres and authors. Thus, all the books were more or less about the pyramids. Second, through an informal survey, the researcher understood that most of the students were interested in topics like the mysteries of Egypt, mummies, and pyramids. Therefore, "Pyramids" was

chosen as the fixed theme for books since they were both available in the market and interesting to the readers. Last, being aware that a combination of unfamiliar vocabularies and lack of interest in one topic would demotivate students, the researcher calculated the difficulty level of the texts by applying both Flesche readability formula and lexical profile. The result was that all texts were fairly easy.

To classify text words into different frequency levels, the lexical profile which was developed by Cobb, updated in 2015, was used. The aim was to categorize words into diverse word families so that a list of words belonging to the 3000-word families could be chosen.

The second pretest which coincided with the VS test was the modified version of Wesche and Paribakht's (1996) VKS test. Unlike the VST, the aim of the VKS test was to assess students' depth of vocabulary knowledge. By using the lexical profile software, four distinct samples of this test were devised, which were inspired by different text words used in the study. Each sample incorporated 20 words belonging to the range of 3000-word families because they were repeated at least five times throughout the predetermined texts. It is said that the more the words are repeated, the better they will be learned (Nation, 2014; Waring & Takaki, 2003; Web, 2007a, 2007b).

After a two-week interval, students were divided into four groups. To avoid test effect, which means examinees might have memorized some of the items from the pretest, some of the scholars recommended a two-week interval (Bardhoshi & Erford, 2017). Group A studied two books in the narrative genre written by one author, while group B studied two books by two different authors in the narrative genre. Then, again, two books written by one author in expository were studied by group C and two books by two authors in expository by group D.

The rationale for choosing the books was first and foremost, to keep the theme fixed to see the effect of NR through different genres and authors, which was the main purpose of this research. Thus, all the books were more

or less about the pyramids. Next, through an informal survey, the researchers understood that most of the students were interested in topics like mysteries of Egypt, mummies, and pyramids. Therefore, “Pyramids” was chosen as the fixed theme for books since they were both available in the market and interesting to the readers.

The number of pages in the narrative genre and the expository genre were not the same. Usually, narrative texts are longer than expository ones. To keep the text length constant, the first 10000 running words beginning from the first word on page one were specified for each group.

At the beginning of each session, the students were encouraged to use their background knowledge and predict what would be discussed on that particular page. Those who had access to expository books benefited from illustrations. On the other hand, students who read narrative texts reviewed what they had read in the previous session and predicted what would happen based on the evidence. Afterward, students were asked to skim the paragraphs for a gist. To avoid creating a threatening atmosphere, the researcher invited the volunteers to discuss their opinions. This helped students to predict the purpose of the passage and understand the main idea. Since the emphasis was on the vocabulary size, the readability level of the

Contrary to traditional English reading classes in Iran, in this class, sentences were not translated line by line. There were two main reasons for such a decision. First, the target group knew most of the high-frequency words; therefore, they could understand approximately 95% of the text. Secondly, students were encouraged to guess the meaning of a word through context clues such as parentheses, colons, and pronouns to become responsible for their own learning rather than being mere consumers. Nevertheless, the only words which were taught explicitly belonged to the range of 3000-word families. The reason was that the students were not exposed to these words repeatedly in contrast to high-frequency words.

Consequently, the probability of learning through implicit instruction was very low.

To compensate for this shortcoming, before starting each session, the researcher applied the lexical profile software to determine the position of the 3000-word families on each page. After students discussed the main idea of each paragraph and became familiar with the topic, the third researcher prompted them to quickly scan the text to find the predetermined words, i.e., 3000-word families. When they did, the third researcher encouraged them to guess the meaning of words. However, guessing the meaning of such words was by far more difficult than high-frequency words because they were used neither in every context nor very frequently. Ryder and Slater (1988) claimed that there is a relationship between word difficulty and its frequency. In other words, the more a word is repeated in the text, the better it can be stabilized in the readers' minds. As a result, the meanings of these words were directly taught both in Persian and English.

There were three explanations for vocabulary direct instruction. First of all, according to Gardner (2008), narrow reading should not be considered as an opportunity to acquire words incidentally, but the words need to be directly taught because of their importance both to the theme in general and the reading text that supports that theme. Second, Stahl (1999) believed that word learning takes a long time and the problem is to use a kind of instruction which can shorten this process. Last, since mid-frequency vocabularies do not occur repeatedly in the context, it is suggested that explicit teaching complement the incidental approach (Cobb, 2007).

At the end of each session, the texts were gathered to avoid checking words at home. To remember the words, it was probable that some of the students used other strategies which might have interfered with the true purpose of NR. To clarify, these actions might upset the equilibrium of the research because it would be unknown whether progress in word learning was the result of NR or other factors.

This procedure lasted for 30 sessions, and each session took an hour. The reason for continuing it for thirty sessions is that vocabulary growth through reading would happen over time (Beck, McKeown, & Kucan, 2002). To conduct the posttest, Wesche, and Paribakht's VKS (1996) test was repeated to measure whether new vocabularies had been learned through narrow reading. As opposed to VKS, VST was not used as a posttest. According to Nation and Beglar (2007), "VST has a different purpose. It is not a diagnostic measure like VLT but is a proficiency measure used to determine how much vocabulary learners know" (p.10). A month later, VKS was administered as a delayed posttest to see how many vocabularies had been retained.

Results

Before running any of these tests, the assumptions underlying a parametric test were checked. For this purpose, skewness, and kurtosis, histograms, and one sample K-S were tested for the normality of distribution and boxplots were used to identify outliers and extremes. To estimate the consistency of VKS scores across two scorers in pretest, posttest, and delayed posttest, inter-rater reliability was calculated. A high degree of reliability was found between the pretest measurements. The average measure of Intra-class correlation coefficients, ICC, in pretest was .971 with a 95% confidence interval from .957 to .981, [F (90, 91) = 35.074, $p < 0.001$]. The average measure ICC in posttest was .987 with a 95% confidence interval from .981 to .992, [F (90, 91) = 77.695, $p < 0.001$]. In delayed posttest, the average measure ICC was .976 with a 95% confidence interval from .967 to .984, [F (90, 91) = 41.989, $p < 0.001$]. It can be concluded that the data were reliable since the results were close to 1.00.

To answer the first research question, i.e., examining the role of genre in reading texts written by OA or MA, i.e., two different genres, two different contexts of authors, and three different times, mixed design factorial ANOVA was applied. Before running it,

descriptive statistics and the underlying assumptions required for applying ANOVA were checked to make sure the assumptions are met.

By observing the means (Table 1), it is obvious that there was a pattern in vocabulary scores of the four groups of participants who read expository or narrative genre written by OA or MA during pretest, posttest, and delayed posttest. The following table shows the descriptive statistics for all the groups.

Table 1.

Descriptive Statistics of Pretest, Posttest and Delayed Posttest Scores of Students Reading Narrative or Expository Texts Written by OA or MA

	Genre	Author	Mean	Std. Deviation	N
WKSPretest	Narrative	OA	35.10	5.531	21
		MA	35.09	6.680	23
		Total	35.09	6.088	44
	Expository	OA	34.81	4.956	21
		MA	31.58	3.624	26
		Total	33.02	4.523	47
	Total	OA	34.95	5.189	42
		MA	33.22	5.516	49
		Total	34.02	5.408	91
WKSPosttest	Narrative	OA	58.57	6.720	21
		MA	49.26	7.671	23
		Total	53.70	8.558	44
	Expository	OA	63.29	6.746	21
		MA	54.58	6.172	26
		Total	58.47	7.723	47
	Total	OA	60.93	7.066	42
		MA	52.08	7.348	49
		Total	56.16	8.438	91
WKSDelayedPosttest	Narrative	OA	48.67	8.039	21
		MA	44.91	8.691	23
		Total	46.70	8.503	44
	Expository	OA	56.81	6.486	21
		MA	48.27	7.598	26
		Total	52.09	8.251	47
	Total	OA	52.74	8.308	42
		MA	46.69	8.219	49
		Total	49.48	8.755	91

Descriptive statistics revealed that all four groups achieved low scores in the pretest. So, there was no main difference between students since the mid-frequency words which belonged to the 3000-word families were either unknown to the students or just familiar. Based on Wesche and Paribakht's (1996) VKS test results, students were awarded 1 or 2 out of four scores which means they did not know the word meaning or they had seen the word but did not remember the meaning respectively. Based on (Table 1) the mean score of GA who was tested on vocabularies of narratives written by OA was 35.10 while the mean score of GB who was about to read narratives written by MA is 35.09. On the contrary, GC who read expository text written by OA, achieved 34.81, whereas GD, who was assigned to read expository texts written by MA, achieved 31.58 in the VKS pretest.

However, the scores of posttests displayed significant growth that had occurred for all four groups. Group A gained 58.57 in contrast to Group B, who scored 49.26. Group C hit the mean score 63.29 in comparison to Group D, who achieved 54.58.

To see whether students have retained the learned vocabularies, delayed posttest scores were examined. Group A achieved 48.67 compared to Group B who gained 44.91. Moreover, the scores of Group C (56.81) differed from Group D scores (48.27).

Table 2.

Tests of Between-Subjects Effects for Genre and Author in VKS Scores

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	591935.274	1	591935.274	6554.942	.000	.987
Genre	591.650	1	591.650	6.552	.012	.070
Author	2118.174	1	2118.174	23.456	.000	.212
Genre * Author	103.280	1	103.280	1.144	.288	.013
Error	7856.419	87	90.304			

From the Tests of Between-Subjects Effects in (Table 2), it can be seen that there was a significant main effect for the ‘Genre’ factor, [$F(1, 87) = 6.552$; $p = .012$], and the effect size was moderate based on Cohen’s (1988) criterion that considers $.01 =$ small effect; $.06 =$ moderate effect; and $.14 =$ large effect ($\eta^2 = .070$).

Therefore, it can be said that there was a significant difference in students’ vocabulary depth scores belonging to either narrative or expository genre, demonstrating that learning mid-frequency words from the expository genre resulted in better performance than reading narrative texts.

The second research question dealt with the role of the author in reading narrative or expository texts. The results of between-subject effects (Table 2) revealed that there was a significant main effect for the ‘Author’ factor, i.e., whether participants read texts written by one or more than one author, [$F(1, 87) = 23.456$; $p = .000$] and the effect size was large ($\eta^2 = .212$). As far as the Author was concerned, reading books written by one author brought about better results than reading books written by more than one author.

To see if there was any significant interaction between genre and author, mixed design factorial ANOVA was used. Based on the results shown in Table 2, there was no significant main effect for the interaction between Genre and Author, [$F(1, 87) = 1.144$; $p = .288$]. Although students who read expository texts or narrative texts written by OA performed better than those groups who read expository or narrative texts written by MA, this interaction was not significant.

The results of the third research question which examined the difference between vocabulary depth scores in pretest, posttest, and delayed posttest were manifested through applying mixed design factorial ANOVA.

Table 3.

Mauchly’s Test of Sphericity for VKS Scores

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^b Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Time	.840	14.971	2	.001	.862	.908	.500

Since there were more than two conditions of the repeated measures variable, i.e., VKS test was repeated three times, the sphericity assumption was checked (Table 3). The Mauchly Test of Sphericity table gives a Mauchly's W test statistic of .840, $df = 2$; $p < 0.05$. Since Mauchly Test of Sphericity is significant, the sphericity assumption is not met. As a result, Greenhouse-Geisser is used since it is an appropriate method correcting the F value. In this case, the value in Greenhouse-Geisser is close to 1.00; hence, one can be confident that issues of sphericity do not affect the calculations (Hinton, 2004).

Table 4.

Test of Within Subjects Effects for VKS Scores

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Squared Eta
Time	Sphericity Assumed	23572.567	2	11786.283	31.870	.000	.859
	Greenhouse-Geisser	23572.567	.724	13669.347	531.870	.000	.859
	Huynh-Feldt	3572.567	1.817	12974.817	531.870	.000	.859
	Lower-bound	23572.567	1.000	23572.567	31.870	.000	.859
Time * Genre	Sphericity Assumed	803.803	2	401.902	18.136	.000	.173
	Greenhouse-Geisser	803.803	1.724	466.112	18.136	.000	.173
	Huynh-Feldt	803.803	1.817	442.429	18.136	.000	.173
	Lower-bound	803.803	1.000	803.803	18.136	.000	.173
Time * Author	Sphericity Assumed	626.755	2	313.377	4.142	.000	.140
	Greenhouse-Geisser	626.755	1.724	363.445	4.142	.000	.140
	Huynh-Feldt	626.755	1.817	344.978	14.142	.000	.140
	Lower-bound	626.755	1.000	626.755	14.142	.000	.140
Time * Genre	Sphericity Assumed	86.759	2	43.379	1.958	.144	.022
	Greenhouse-Geisser	86.759	1.724	50.310	1.958	.151	.022
	Huynh-Feldt	86.759	1.817	47.754	1.958	.149	.022
	Lower-bound	86.759	1.000	86.759	1.958	.165	.022
Error (Time)	Sphericity Assumed	3855.857	174	22.160			
	Greenhouse-Geisser	3855.857	150.030	25.701			
	Huynh-Feldt	3855.857	158.061	24.395			
	Lower-bound	3855.857	87.000	44.320			

By looking at Greenhouse-Geisser rows (Table 4), it can be seen that there was a main effect for the repeated measure factor ' Time ', [F (1.724, 150.030) = 531.870; $p < 0.01$] with a very large effect size (eta squared = .859). There was also a significant main effect for the interaction of 'Time and Genre', [F (1.724, 150.030) = 18.136; $p < 0.01$] and very large (eta squared = .173) main effect size. Furthermore, there was a main effect for the interaction of ' Time and Author ', [F (1.724, 150.030) = 14.142; $p < 0.01$] while the main effect was strong (eta squared =.140). However, there was not a main effect for the interaction of 'Time, Genre and Author', [F (1.724, 150.030) = 1.958; $p = 0.151$] and the effect size was small (eta squared =.022).

Comparison of the scores attained from three phases i.e., pretest, posttest, and delayed posttest show that students had a different performance in each of these phases. To understand where this difference lay, pairwise comparison was conducted. Considering the interaction between the VKS scores attained through three different times and Genre, there were differences between the pretest, posttest, and delayed posttest scores of those who read the narrative or the expository genre. Students' scores were higher when reading the expository genre in both the learning and retention phases. Also, there was a significant interaction between the VKS scores and the Author. When students read books written by one author, they achieved higher scores in both the learning and retention phases. On the other hand, reading books written by more than one author did not result in higher scores. However, there was not a significant interaction between the three factors of the genre, author, and VKS scores. Therefore, a combination of genre and the number of authors did not make a difference among scores in three VKS administrations.

Table 5.

The Mean and Standard Error of Scores across Three VKS Phases

Time	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	34.142	.554	33.041	35.244
2	56.424	.719	54.995	57.853
3	49.665	.817	48.042	51.288

Table 5 indicates the mean of scores that students achieved across the three phases. Hence, scores that students achieved in each of VKS administrations are compared to see where the difference lies. The pairwise comparison table gave us a comparison of means for all paired combinations of the repeated measures factor, which was the VKS test. All comparisons were adjusted using the Bonferroni method.

Table 6.

Pairwise Comparison of VKS Factor Comparing Students Depth Scores in Three Times

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
					Lower Bound	Upper Bound
1	2	-22.282*	.705	.000	-24.002	-20.561
	3	-15.522*	.811	.000	-17.502	-13.543
2	1	22.282*	.705	.000	20.561	24.002
	3	6.759*	.564	.000	5.383	8.136
3	1	15.522*	.811	.000	13.543	17.502
	2	-6.759*	.564	.000	-8.136	-5.383

Table 6 shows all the possible comparisons for the three levels of repeated measures variable, i.e., Time. As can be seen from the table, all the possible pairwise comparisons were significant since all the values were less than 0.01.

Table 7.

Pairwise Comparison between Three Phases of VKS Test and Expository vs. Narrative Genre

Time	(I) Genre	(J) Genre	Mean Difference (I-J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
						Lower Bound	Upper Bound
1	Narrative	Expository	1.898	1.108	.090	-.305	4.101
	Expository	Narrative	-1.898	1.108	.090	-4.101	.305
2	Narrative	Expository	-5.015*	1.438	.001	-7.873	-2.157
	Expository	Narrative	5.015*	1.438	.001	2.157	7.873
3	Narrative	Expository	-5.750*	1.633	.001	-8.996	-2.503
	Expository	Narrative	5.750*	1.633	.001	2.503	8.996

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Since there was a significant interaction between Time and Genre (Table 4), the simple main effect was assessed to examine the effect of 'Genre' at each level of Time separately. According to Table 7, in Time 1 or pretest, the difference (between narrative and expository groups) was not significant ($p=.090$). However, considering the second block of the table, Time 2 or posttest, it can be seen that readers performed significantly differently by reading the two different genres ($p=.001$). In addition, in Time 3 or delayed posttest, the difference was also significant ($p = .001$).

Table 8.

Pairwise Comparison between Three Phases of VKS Test and OA vs. MA Texts

Time	(I) Author	(J) Author	Mean Difference (I-J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
						Lower Bound	Upper Bound
1	OA	MA	1.620	1.108	.147	-.583	3.824
	MA	OA	-1.620	1.108	.147	-3.824	.583
2	OA	MA	9.010*	1.438	.000	6.152	11.867
	MA	OA	-9.010*	1.438	.000	-11.867	-6.152
3	OA	MA	6.147*	1.633	.000	2.901	9.393
	MA	OA	-6.147*	1.633	.000	-9.393	-2.901

Since there was a significant interaction between Time and Author (Table 4), the simple main effect was assessed to examine the effect of 'Author' at each level of Time separately. In Time 1 or pretest (Table 8), the difference (between OA and MA groups) was not significant ($p=.147$). However, considering Time 2, it can be seen that readers performed significantly differently by reading the texts written by OA or MA ($p< 0.05$). In addition, in Time 3 or delayed posttest, the difference was also significant as $p<0.05$.

By comparing vocabulary scores attained through reading narrative vs. expository texts during pretest, posttest, and delayed posttest, it can be said that there was no difference between scores before reading such texts. However, differences appeared when some of the students read narrative texts, and others read expository texts. In fact, those students who read expository texts outperformed their counterparts, i.e., narrative readers. Even during delayed posttest, expository readers retained more mid-frequency vocabularies in contrast to those who read narrative texts.

It is also observed that there was no difference between students' depth scores during the pretest; however, as students started to read OA texts, they overtook MA text readers in the posttest. In addition, students who read OA texts were more successful in retaining vocabularies during delayed posttest.

It is also observed that the group who read narratives written by one author outperformed the other group who read narratives written by more than one author in both learning and retention phases. Comparing the mean differences of scores in the learning and retention phase, it is obvious that there was less attrition in the knowledge of the students who read narrative texts written by more than one author. Comparison of narrative vs. expository texts written by OA vs. MA shows that while all groups performed significantly in the posttest phase, reading expository texts either written by OA or MA resulted in better vocabulary retention after a month interval.

Discussion

This study was conducted in an attempt to clarify whether NR, that is, reading books written on a particular genre (expository vs. narrative), author (one vs. more than one) and theme (pyramids) could have any effect on learning mid-frequency words especially the 3000-word families.

The first null hypothesis that was “*Genre does not play a significant role in reading OA or MA texts*” was rejected. That is, regardless of the fact that texts were written by OA or MA, there was a significant difference between expository and narrative texts, while a fixed theme was predetermined. In other words, expository readers performed better than narrative readers. Thus, it is in line with Gardner’s (2008) hypothesis that “themes have their greatest impact on specialized vocabulary recycling among informational (expository) materials” (p. 108).

The reason why the presence of a particular theme makes a difference in terms of vocabulary learning among expository texts rather than narrative ones is rooted in the nature of these texts. As opposed to narrative texts which narrate a series of events, the main purpose of expository texts is to provide information about a particular theme. Therefore, reading books on a particular theme and repetition of thematically related words make a meaningful connection between words in the minds of readers. This result is in line with Ausubel’s (1968) theory that readers will subsume the meaning of related words because they are stored through association.

The second null hypothesis “*number of authors does not play any significant role in both narrative and expository texts*” was rejected. In other words, by comparing OA vs. MA texts in both narrative and expository texts, it was demonstrated that those groups who read OA texts achieved higher scores than those reading MA texts.

The results of this research question confirmed Ahmadi and Poordaryaie’s (2015) research. They claimed that narrowing texts to the author, i.e., reading texts written by one author, would enable readers to

develop more vocabulary. In this study, although the effect of theme rather than the author was more observable in expository texts, it could not be ignored that there was a difference between the group who read expository books written by MA and the group who read expository books by OA. In addition, the same condition happened for students who read narratives written by OA or MA.

Thus, this finding, too, supported Gardner's (2008) view, who believed in authorship as a determining factor in narrative texts. "Authentic children's narratives written by the same author have substantially more specialized vocabulary recycling than narratives written by several different author" (p.108).

However, the findings from this research question did not support Gardner about the fact that "authorship has no observable impact on specialized vocabulary recycling among authentic children's expository materials" (p. 108). Texts with one author appeared to make a difference among narrative texts, and so did it with regard to expository texts.

Analysis of expository texts written by OA reveals that the percentage of author-based words which not only belong to the range of 3000-word families but also were repeated frequently was 8.06% on average. On the other hand, MA expository texts written about the same theme had only 5.85% of words in common. While the percentage of the author-based words was too minimal among OA expository texts, it was sufficient to make a significant difference between students' scores that read OA or MA expository texts.

The third null hypothesis, i.e., "*there is no significant interaction between genre and author*" was accepted. According to inferential statistics, although each factor was influential by itself, no interaction was found between them. The reason may lie in the fact that a fixed theme was determined for all groups. Investigations of the texts highlighted some important points about the genre, author, and theme in relation to mid-frequency words. First, while the percentage of the author based mid-frequency words in OA narratives was

more than other groups, the number of such words was not very large. In fact, in the case of narratives, the author in companion with the theme provided a condition for the repetition of mid-frequency words. On the other hand, while expository texts were under the influence of theme for the recycling mid-frequency words, the number of authors i.e., reading single-author texts also facilitated this process.

The last null hypothesis, "*there is no significant difference between vocabulary depth scores in pretest, posttest, and delayed posttest* " was rejected. The second question dealt with the progress of students who read narrowly to examine how many words they have learned and retained. While vocabularies belonged to the 3000-word families were either unknown or just familiar to the students during the pretest, the posttest scores confirmed the efficacy of narrow reading because of the significant improvement of all of the students. However, results achieved from the delayed posttest were different. While there was less attrition in the knowledge of the two groups reading expository texts, most of the mid-frequency vocabularies learned through reading OA narratives sunk into oblivion during the retention phase. Nevertheless, the scores of students who read MA narratives neither increased significantly in the learning phase nor reduced drastically during the delayed posttest. It is probable that readers of MA texts forget most of the words after the first exposure because learning these words was not reinforced by another meeting (Nation, 1997).

One of the reasons which can explain why the group who read OA narratives forgot mid-frequency words faster than other groups is the high percentage of the author based mid-frequency words. Contrary to the thematic set recommended by Tinkham (1997), author-based words do not share similar schema with previously relevant entities in mind since they are frequent just in the texts written by a particular author. Consequently, there is no chance to make a meaningful connection with the words under the same category, and such words will be stored in a rote manner. The result is that

“their retention is influenced primarily by the interfering effects of similar rote materials learned immediately before or after learning task” (Brown, 2007, p. 94) and are forgotten very soon.

Some of the findings of this study support Abdollahi and Farvardin’s (2016) opinion. They believed that narrow reading provides a great opportunity for the development of learner’s vocabulary recall and retention. Abdollahi and Farvardin claimed that repetition of words in texts narrowed to a particular theme led to word consolidation in learners’ minds. Still, they did not consider the frequency of words as an element which affected the process of word exposures. Usually, there is a small number of words that recur very frequently in every text while a large number of words are part of mid or low-frequency words (Schmitt & Schmitt, 2014). Therefore, it is possible that students read thematically related texts, but they would not encounter words repeatedly because these words are less frequent.

This study also confirms the previous research that emphasized the effect of word repetition on word retention. According to Laufer and Rozovski (2011), it is not definite how many word encounters is enough. However, the posttest and delayed posttest results provide evidence for the necessity of word encounters. That is, the more they are exposed to mid-frequency words, the better they will retain these words. Contrary to group B, who read narratives written by more than one author, other groups were exposed to 3000-word families more frequently. Consequently, they remembered more words after a month interval.

The results of this study may also provide reasons for the superiority of NR over ER in the case of teaching or learning mid-frequency words. First and foremost, since mid-frequency words do not occur frequently in every context (Cobb, 2007), the ER would not be an appropriate approach to learn such words. The reason is that readers will be bombarded with a large number of words rather than having the chance to meet these words again; hence, words will be forgotten very soon. Second, as opposed to high-frequency

words, mid-frequency words need to be learned incidentally as well as explicitly; nevertheless, such a need cannot be provided by ER since the main focus of ER is on understanding the message of texts while learning vocabulary implicitly. On the other hand, NR is not only an opportunity for incidental vocabulary learning but also teaching vocabulary directly (Gardner, 2008). Last but not least, the main purpose of the ER is deriving pleasure from reading. Therefore, students are provided with manipulated books such as graded readers to be able to guess the word meanings through the context. However, unassisted comprehension of authentic texts, which incorporates not only high-frequency words but also mid and low-frequency ones, requires 98% coverage of a text (Nation, 2006). It seems that NR is far preferable to the ER. When texts are narrowed to a particular author, genre, or theme, the chance of vocabulary occurrence from different frequencies will be increased. Therefore, repeated exposure, as well as direct teaching, stabilize words especially mid-frequency ones in the readers' minds.

In conclusion, although each of NR elements like genre or author played a significant role in learning and retention of mid-frequency words, the effect of the theme cannot be ignored. Text analysis (Table 1) revealed that most of the words which happened in the texts were thematically related. Therefore, the theme in relation to author or genre can be a powerful element which can expose students to mid-frequency words repeatedly.

Conclusion and Implications

The purpose of this study was to see whether Narrow Reading which means reading in only one genre, one subject matter, or the work of one author, would lead to learning and retention of the mid-frequency words. The results revealed that those who read expository texts outperformed those who read the narrative in learning and retention of mid-frequency words. Likewise, reading texts written by OA was advantageous in reading MA texts. However, just as reading OA narrative texts was effective, so was reading OA expository

texts. It implies the fact that each of these factors was influential by itself rather than in combination with each other. Thus, it is suggested that researchers or teachers, who want to conduct a similar study, assign students to different types of NR. On the basis of the results of this study, one may pose the superiority of NR to other reading approaches, e.g., extensive reading, in the case of learning mid-frequency words. The first reason is that students will have the chance to meet the words repeatedly, so they will remember the words for a longer time. Secondly, since mid-frequency words do not occur frequently in every context, NR provides an opportunity for incidental as well as intentional vocabulary learning which may stabilize words in the minds of readers. Last, NR helps readers to give up the belief that they do not know mid-frequency words in authentic texts since they are provided with the opportunity to meet the word repeatedly. In fact, the more readers do NR, the better they can comprehend unfamiliar words.

Considering the implications, it seems that one of the groups which benefits enormously from this research is university students since most of these students need to know technical words to be able to read specialized texts. However, they do not have the chance to meet these words in every context since they are not very frequent. Thus, NR sounds to be a good solution. When students read thematically related texts, it is recommended to narrow texts to expository genre or texts written by one author so that they will have the chance to meet mid-frequency words repeatedly.

Another group that may take advantage of NR is those who are interested in learning vocabulary. The problem is that such texts are not controlled for vocabularies; therefore, students would encounter a wide range of vocabularies from different frequency levels. However, when students read a series of books written by a particular author and about a particular theme, they will be exposed to these words repeatedly since each author has a set of favorable phrases and expressions which are frequent in his works. Furthermore, providing students with a series of stories by a favorite author

may help readers give up the belief that reading the works of a particular author is tiring and tedious.

The next application of this research is related to teachers because they will have the opportunity to present mid-frequency words in classrooms. It is believed that teachers do not deal with mid-frequency words to a great extent (Schmitt & Schmitt, 2014) because these words cover a small percentage of running words in different texts. However, on the basis of this study, students will be encouraged to read authentic texts narrowly. Authentic texts are not controlled for the presentation of vocabulary. It means that students will be exposed to high, mid, and low-frequency words. Therefore, when students read texts narrowly, they would encounter different ranges of words repeatedly so that they would give up the belief that authentic texts are not comprehensible.

Still, the most important application of this study is that Gardner's (2008) theory has been put into practice. He theorized that expository texts are under the influence of theme to recycle specialized words while the number of authors makes a difference in narrative texts. The results of this study substantiated the fact that while theme plays a significant role in expository texts, the number of authors is also important. The same thing happened for narratives. While stories written by a particular author expose students to mid-frequency words repeatedly, the presence of a particular theme complements this process. It should be mentioned to cross-examine Gardner's theory, the role of the same two genres, i.e., expository and narrative, were controlled and other kinds of genre were not investigated which could be considered as the delimitation of this study and could be considered as the area for further research in addition to considering other factors such as different majors especially English ones or different ages and genders.

References

- Abdollahi, M., & Farvardin, M. T. (2016). Demystifying the effect of narrow reading on EFL learners' vocabulary recall and retention. *Education Research International*, 2016. Article ID 5454031, 1-10. DOI: [10.1155/2016/5454031](https://doi.org/10.1155/2016/5454031)
- Ahmadi, A., & Poordaryaie, N. (2015). Evaluating the effectiveness of narrow reading on vocabulary development among EFL learners in Bastak town. *The Indian Journal of Fundamental and Applied Life Sciences*, 5(S4), 844-864.
- Alderson, J.C., & Urquhart, A.H. (1984). Postscript on Hosenfeld. In J.C Alderson & A.H., Urquhart (Eds). *Reading in a Foreign Language* (pp. 245-249). London: Longman.
- Ausubel, D. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart & Winston.
- Bardhoshi, G., & Erford, B. T. (2017). Processes and procedures for estimating score reliability and precision. *Measurement and Evaluation in Counseling and Development*, 50(4), 256-263.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction*. NY: Guilford Press.
- Brown, H. D. (2007). *Principles of language learning and teaching*. NY: Pearson Education, Inc.
- Brown, J. (2009). *The great Egyptian grave robbery: Flat Stanley's Worldwide Adventures Series*. New York, NY: HarperCollins.
- Cho, K. S., Ahn, K. O., & Krashen, S. (2005). The effects of narrow reading of authentic texts on interest and reading ability in English as a foreign language. *Reading Improvement*, 42(1), 58- 64.
- Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning & Technology*, 11(3), 38-63.

- Elley, W. B. (1991). Acquiring literacy in a second language: The effect of book-based programs. *Language learning*, 41(3), 375-411. Doi: 10.1111/j.1467-1770.1991.tb00611.x
- Filer, J. (2003). *The mystery of Egyptian mummy*. London: British Museum Press.
- Filer, J. (2005). *Pyramids*. United States of America: Oxford University Press.
- Folse, K. (2010). Is explicit vocabulary focus the reading teacher's job? *Reading in a Foreign Language*, 22(1), 139- 160. Retrieved from <https://files.eric.ed.gov/fulltext/EJ887885.pdf>
- Gardner, D. (2008). Vocabulary recycling in children's authentic reading materials: A corpus-based investigation of narrow reading. *Reading in a Foreign Language*, 20(1), 92-122.
Retrieved from <http://files.eric.ed.gov/fulltext/EJ793218.pdf>
- Gorman, A. M. (1961). Recognition memory for nouns as a function of abstractness and frequency. *Journal of experimental psychology*, 61(1), 23-29.
- Hafiz, F. M., & Tudor, I. (1989). Extensive reading and the development of language skills. *ELT Journal*, 43(1), 4-13. doi:10.1093/elt/43.1.4
- Hall, J. F. (1954). Learning as a function of word-frequency. *The American journal of psychology*, 67(1), 138-140.
- Hansen, L. E., & Collins, P. (2015). Revisiting the case for narrow reading with English language learners. *The Reading Matrix: An International Online Journal*, 15(2). Retrieved from <http://www.readingmatrix.com/files/13-ro1o846k.pdf>
- Hatch, E., & Farhady, H. (2002). *Research design and statistics*. Tehran, Tehran: Rahnama Publications.
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11, 207-223.
Retrieved from <http://nflrc.hawaii.edu/rfl/PastIssues/rfl112horst.pdf>

- Horst, M., Collins, L., & Cardoso, W. (2009). Focus on vocabulary in ESL teacher talk. In the *annual conference of the American Association for Applied Linguistics*. Denver, Colorado.
- Hunt, A., & Beglar, D. (2005). A framework for developing EFL reading vocabulary. *Reading in a Foreign Language* 17(1), 23-59.
- Jenkins, J. R., M.L. Stein & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal*, 21,767-787. DOI:10.3102/00028312021004767
- Kang, E. Y. (2015). Promoting L2 vocabulary learning through narrow reading. *RELC Journal*, 46(2), 165-179. DOI:10.1177/0033688215586236
- Krashen, S. (1981). The case for narrow reading. *TESOL Newsletter*, 15, 23-25.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, 73(4), 440-464.
- Krashen, S. (2004). The case for narrow reading. *Language Magazine*, 3(5), 17-19. Retrieved from <http://www.sdkrashen.com/content/articles/narrow>
- Kyongho, H., & Nation, P. (1989). Reducing the vocabulary load and encouraging vocabulary learning through reading newspapers. *Reading in a Foreign Language*, 6(1), 323-335. Retrieved from <https://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/1989-Hwang-Newspapers.pdf>
- Laufer, B. (1996). The lexical threshold of second language reading comprehension: What it is and how it relates and how it relates to L1 reading ability. In K. Saja-Vaara, & C. Fairweather (Eds.), *Approaches to Second Language Acquisition* (pp. 55-62). Jyvaskyla, Finland: University of Jyvaskyla.

- Laufer, B., & Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence, and their combination. *Language Teaching Research*, 15(4), 391-411. Retrieved from <http://journals.sagepub.com/doi/pdf/10.1177/1362168811412019>
- Machen, A. (2010). *The shining pyramid*. Whitefish, Montana: Kessinger Publishing.
- Mackey, A., & Gass, S.M. (2005). *Second language research: Methodology and design*. London: Lawrence Erlbaum.
- Meara, P. M., & Milton, J. (2003). *X_Lex: The Swansea vocabulary levels test*. Newbury: Express Publishing.
- Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American educational research journal*, 24(2), 237-270. <https://doi.org/10.3102/00028312024002237>
- Nation, I. S. P. (1982). Beginning to learn foreign vocabulary: A review of the research. *RELC Journal*, 13, 14-36. DOI: 10.1177/003368828201300102
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. United Kingdom: Cambridge United Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59–82. Retrieved from https://www.lexutor.ca/cover/papers/nation_2006.pdf
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The language teacher*, 31(7), 9-13.
- Nation, P. (1997). The language learning benefits of extensive reading. *LANGUAGE TEACHER-KYOTO-JALT-*, 21(5), 13-16. Retrieved from <http://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/1997-Benefits-of-ER>

- Nation, P. (2012). The vocabulary size test. Retrieved from <https://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Vocabulary-Size-Test-information-and-specifications.pdf>
- Nation, P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language*, 26(2), 1. Retrieved from <http://nflrc.hawaii.edu/rfl/October2014/articles/nation.pdf>
- Paya-Guerrero, M., & Segura, M. L. E. (2015). EFL vocabulary acquisition: Narrow reading versus reading plus vocabulary-enhancement activities. A case study with Spanish secondary school students. *Porta Linguarum: revista internacional de didáctica de las lenguas extranjeras*, 24, 93-105. Retrieved from http://www.ugr.es/~portalin/articulos/PL_numero24/7MARAVILLAS.
- Perry, D., & MacDonald, P. (2001). Word knowledge and vocabulary instruction. Paper presented in the *International conference on engineering education*, 6-10 August, Oslo, Norway.
- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18, 1-28. Retrieved from <http://nflrc.hawaii.edu/rfl/April2006/pigada/pigada.pdf>
- Saragi, T., Nation, I. S. P., & Meister, G. F. (1978). Vocabulary learning and reading. *System*, 6(2), 72-78. DOI: 10.1016/0346-251X(78)90027-1
- Sataryah, C., Nation, I. S. P., & Kennedy, G. (1994). How useful is EAP vocabulary for ESP? *RELC Journal*, 25(2), 34-50. DOI: 10.1177/003368829402500203
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. London: Palgrave Macmillan.
- Schmitt, N., & Carter, R. (2000). The lexical advantages of narrow reading for second language learners. *TESOL Journal*, 9(1), 4-9. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.707.3992&rep=rep1&type=pdf>

- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(04), 484-503. Retrieved from http://eprints.nottingham.ac.uk/32297/1/A_Reassessment_of_Frequency_and_Vocabulary_Size_in_Vocabulary_Teaching_-_Final.pdf
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55-88.
- Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York, NY: Guilford Press.
- Stine, R. L. (1993). *The curse of the mummy's tomb*. Sydney, New South Wales: Scholastic Australia.
- Stine, R. L. (2006). *Return of the mummy*. Logan, Utah: Perfection Learning Corporation.
- Taplin, S. (2003). *Mummies & pyramids*. New York, NY: Scholastic INC.
- Tinkham, T. (1997). The effects of semantic and thematic clustering on the learning of second language vocabulary. *Second language research*, 13(2), 138-163. DOI:10.1191/026765897672376469
- Waring, R. (2002). Scales of vocabulary knowledge in second language vocabulary assessment. *Kiyo, The Occasional Papers of Notre Dame Seishin University*, 46(1), 35-41.
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15(2), 130-163.
- Webb, S. (2005). Receptive and productive vocabulary learning: The effect of reading and writing on word knowledge. *Studies in Second Language Acquisition*, 27 (1), 33-52. DOI:10.1017/S0272263105050023
- Webb, S. (2007a). Learning word pairs and glossed sentences: The effects of a single context on vocabulary knowledge. *Language Teaching Research*, 11(1), 63-81. DOI:10.1177/1362168806072463

- Webb, S., & Rodgers, M. P. (2009). Vocabulary demands of television programs. *Language Learning*, 59(2), 335-366. DOI: 10.1111/j.1467-9922.2009.00509.x
- Webb, S., & Rodgers, M. P. (2009). The lexical coverage of movies. *Applied Linguistics*, 30(3), 407-427. DOI:10.1093/applin/amp010
- Wesche, M., & Paribakht, T. S. (1996). Assessing Second Language Vocabulary Knowledge: Depth versus Breadth. *Canadian Modern Language Review*, 53(1), 13-40. Retrieved from <https://eric.ed.gov/?id=EJ535872>
- Woods, M., & Woods, M. B. (2008). *The tomb of King Tutankhamen*. Minneapolis, Minnesota: Twenty-First Century books.
- Yang, A. (2001). Reading and the non-academic learner: a mystery solved. *System*, 29(4), 451-466. DOI: 10.1016/S0346-251X (01)00038-0

Appendix A.**List of 20 Words in the Narrative Text by One Author**

- | | | | | |
|------------|--------------|-------------|-------------|-------------|
| 1. ancient | 2. beam | 3. chamber | 4. Curve | 5. gaze |
| 6. nod | 7. sigh | 8. trail | 9. treasure | 10. whisper |
| 11. Echo | 12. entrance | 13. explore | 14. Glow | 15. grip |
| 16. pit | 17. priest | 18. Silent | 19. squeeze | 20. violent |

Appendix B.**List of 20 Words in the Narrative Text by Two Authors**

- | | | | | |
|----------------|------------|------------|-------------|-------------|
| 1. conclusion | 2. confess | 3. device | 4. evident | 5. hint |
| 6. puzzle | 7. theory | 8. tragedy | 9. vague | 10. whisper |
| 11. archeology | 12. crew | 13. disc | 14. nod | 15. sigh |
| 16. treasure | 17. vast | 18. sweat | 19. squeeze | 20. urgent |

Appendix C.**List of 20 Words in the Expository Text by one Author.**

- | | | | | |
|---------------|---------------|--------------|--------------|------------|
| 1. archeology | 2. carve | 3. ceremony | 4. funeral | 5. goods |
| 6. grave | 7. hip | 8. internal | 9. layer | 10. organ |
| 11. preserve | 12. Priest | 13. religion | 14. scan | 15. wealth |
| 16. pit | 17. Architect | 18. Mine | 19. entrance | 20. cattle |

Appendix D.**List of 20 Words in the Expository Text by Two Authors**

- | | | | | |
|---------------|---------------|--------------|----------------|--------------|
| 1. archeology | 2. carve | 3. entrance | 4. exhibit | 5. funeral |
| 6. organ | 7. permission | 8. preserve | 9. incorporate | 10. religion |
| 11. reveal | 12. scan | 13. treasure | 14. wealth | 15. chamber |
| 16. goods | 17. Grave | 18. portrait | 19. priest | 20. palace |