

# An Investigation into the Effects of Topic Familiarity on Comprehension and Vocabulary Learning

Burcu Varol

Department of English Language Teaching, Yildiz Technical University, Davutpasa, Istanbul, Turkey

E-mail: [bvarol@yildiz.edu.tr](mailto:bvarol@yildiz.edu.tr)

## Abstract

This study aims to understand the relationship between text comprehension and incidental vocabulary gain, and the effect of topic familiarity on these variables. Freshman students of the ELT department at a state university took part in this repeated-measures design. A background questionnaire was employed to decide more and less familiar topics for the present group. On a five-point Likert scale, they were asked to rate how much familiar they are with the listed topics. Following the administration of the questionnaire, they read four contrived script-based texts two of which have a more familiar content, while the remaining two texts present less familiar information. Upon reading, they completed a free recall task. Passage sight vocabulary was tested 2 days and 18 days after reading the texts through vocabulary recognition and vocabulary production tasks. Correlational analyses and t-test results depicted robust effects of topic familiarity for all variables; however, the relationship between text comprehension and vocabulary gains was mediocre.

**Keywords:** Topic familiarity; incidental vocabulary acquisition; reading in a second language

## 1. Introduction

As a means to accessing information sources, reading is a vital language skill whose gradual development requires intensive and extensive exposure to print. Basically, reading “entails converting print into language and then to the message intended by the author” (Koda, 2007; p. 1). Within the realm of the interactive models in reading, the reader is no longer a passive decoder of textual input, s/he is rather an active participant in the reading process who rebuilds textual representation encoded by the writer based upon his/her prior knowledge (Anderson, 1999; Maria, 1990). Fluent reading enables readers to spare their resources for other processes as they read fast and accurately and make use of their knowledge of syntax and prosody automatically without paying conscious attention to these forms (Grabe, 2004). Yet, in L2 reading, although readers have already developed literacy skills in their L1, other linguistic constraints such as limited vocabulary or grammar knowledge hamper comprehension and the smooth flow of reading. Given the vital role topic familiarity and word knowledge play in comprehension, this study through a repeated-measures design attempts to enrich our understanding of how these factors assist the comprehension process.

### 1.1 Literature review

#### 1.1.1 Models of reading comprehension

Grabe (1991) describes reading as a fluent, purposeful, interactive, and flexible process that aggregates step by step whose ultimate aim is comprehension. According to the process theories, reading is a sequential, bottom-up decoding process (Gough, 1972; as cited in Urquhart & Weir, 1998). At odds with this view, Goodman (1967) proposed a top-down approach to reading where it is described as extracting meaning through the constant predictions of the readers which are confirmed or rejected in the light of the readers’ background knowledge. In the interactive models of comprehension, however, information processing is viewed as the simultaneous processing of lexical, semantic, and syntactic input along with the previous experiences of readers organized around schemata (Rumelhart, 1977; Stanovich, 1980). Based upon Rumelhart’s (1977) interactive model to reading, Eskey (1988) advances a combination of bottom-up and top-down models wherein the good reader is the one who is both a decoder and a good interpreter employing every information source equally to reconstruct the message.

Building on the idea that “comprehension process can be decomposed into components” (p. 364), Kintsch and van Dijk (1978) proposed a model for understanding processes underlying text comprehension and for analyzing recall and summary protocols. It was maintained in their construction-integration (CI) model that knowledge consists of associative networks of concepts and textual propositions, and global knowledge structures in the form of schemas and frames all contribute to mental model development (Kintsch, 1988). The first stage of mental model building is decoding the surface structure which takes place at the linguistic level of processing words and phrases, and it is represented as a set of propositions. Constructing text base which is “a hierarchical propositional representation of the information within the text” (Salmerón, Kintsch & Cañas, 2006, p. 1158) is another level of comprehension. Text base or semantic structure contains microstructure and macrostructure levels. Microstructure is formed at the local level with the integration of textual propositions and inferences in a coherent and related network. Macrostructure contains the global structure of the text and formed

by synthesizing microstructures into higher-order units. Situation model is, on the other hand, built with the integration of textual information formed in the text base into the reader's already existing knowledge prior to reading.

#### *1.1.2 The role of word knowledge in reading*

In Hudson's (2007) operationalization, reading is seen as "a complex activity that involves combination of factors, such as grapheme recognition, phonological representation (perhaps), syntactic structure, background knowledge, processing strategies, text structure, understanding vocabulary (mixed with background knowledge?), and the context of reading act" (p. 289). All of these components are accepted as prerequisites for fluent and successful second language reading (Grabe, 1991; Koda, 2007). Reading in a second language is generally slow probably due to a lack of automaticity in word recognition (Anderson, 1999) as a result of which the reciprocal relationship between vocabulary gain and reading comprehension is lost leading to limited exposure to texts (Nuttall, 2005). It has been contended that poor L2 readers spend so much time on word by word decoding that it gets extremely difficult to process and store textual information in their working memory. Hence, it leads to poor comprehension due to a lack of word-recognition automaticity (Bernhardt, 1991), and this in turn restricts the extensive reading experiences. The place of vocabulary knowledge in reading research has always been central (Bernhardt, 1984) and word knowledge has been consistently found to be an important factor in reading comprehension (Bernhardt & Kamil, 1995; Davis, 1968; Nation, 2001). Concerning the relationship between reading comprehension and vocabulary, Sinatra, Brown and Reynolds (2002) put forward that good readers do not devote their attention to such strategies as guessing the meanings of words as they already have a good repertoire of vocabulary knowledge that enables quick word identification for these readers. Instead, they save their attentional resources for conscious and intentional processing. It follows that a large mental lexicon is a prerequisite for effortless comprehension.

Whether or not vocabulary instruction should be isolated or incidental has been a major area of debate in the literature. In view of the reciprocal relationship between reading and incidental vocabulary learning, research has established that learning to read in a second language is also learning new words while the development of mental lexicon at the same time enhances reading comprehension (Grabe & Stoller, 1997). Furthermore, comprehension-based approaches to reading model the steps of first language acquisition in explaining learning through reading (Paribakht & Wesche, 1993). Just as listening precedes speaking in first language acquisition and as ability to compose texts develops subsequent to the ability to read fluently, second language acquisition emerges through comprehension as readers are exposed to and learn from the texts they process. It is contended that when the focus of learners is on language during meaning extraction, this exposure would lead to the gradual internalization of second language syntax and vocabulary.

Krashen (1989) claimed that comprehensible input is enough on its own for lexical intake; however, differing results emerged from the studies as to the effects of incidental learning. As a case in point, Jacobs, Dufon, and Fong (1994) explored the impact of glossing on incidental learning of words through reading. To do so, they asked the participants to read an expository text and note down as much as they remembered from the text. For the assessment of vocabulary gains, they were required to provide L2 to L1 translations of the targeted words in the text. The findings showed a modest to strong correlation between the lexical gains and text comprehension for the participants who read the text with the glosses. Nevertheless, this incidental learning did not turn into durable vocabulary learning as the effects of glossing disappeared after 4 weeks. On the other hand, incidental learning of words was seen to be converted into long-term learning in some other studies. Rott (1999), for instance, investigated the relationship between text comprehension and vocabulary gains and retention of intermediate learners of German. By adopting L2 to L1 translation and multiple-choice translation tasks as indicators of receptive vocabulary, lexical gain and retention scores of the participants were measured. In the end, a moderate to strong correlation was noted which was also observed to be strengthened over time.

Although vocabulary learning has been proven to take place through incidental learning during reading (Horst, Cobb & Meara, 1998; Paribakht & Wesche, 1997), it requires massive amounts of extensive reading and repeated exposure to the newly learned lexical items in different contexts (Koda, 2005). For instance, Nagy, Herman, and Anderson's (1985) study depicted that the chance of eliciting the correct definition of words encountered in a text is only one in ten for school-age children in an L1 context. In addition, in a very large-scale study, Horst et al. (1998) revealed that despite multiple exposures to the target words, the participants in their study were able to identify only five words among the thousands of words in the story. Consequently, they claim that for incidental learning to take place, readers need to allocate considerable time and effort to word learning since reading in L2 is already a slow and laborious process.

#### *1.1.3 The role of topic familiarity in reading*

There is considerable evidence that the higher the levels of background knowledge are, the richer the textual analyses and interpretations corroborated by the availability of schemata in cognition will be (Alptekin & Erçetin, 2011; Leaser, 2007; Ozuru, Dempsey & McNamara, 2009). The contribution of background knowledge in the meaning-making process is called the schema theory (Carrell & Eisterhold, 1988). According to this theory, it is

the reader who initiates the meaning construction process rather than an over-reliance on text. Meaning is formed as a result of the interaction between the text and the reader's background knowledge. To be able to restore meaning, the reader should have relevant prior knowledge about the written material, thus the reader with the more background knowledge has more chance of comprehending the text. In line with this approach, Connectionist Principles of comprehension (e.g., St. John & McClelland, 1992) view knowledge emerging during reading firstly by constructing text-base upon employing bottom-up processing of textual input which is followed by integrating it into reader's global knowledge. Therefore, the role of background knowledge is seen central in the reading comprehension process.

Coady (1979) depicts a model of the ESL reader and places background knowledge at the heart of it. Accordingly, the comprehension process is an amalgamation of a) conceptual abilities which are described as the intellectual abilities, such as the ability to analyze, synthesize and make inferences, b) process strategies which refer to the mental processes employed for different purposes, and c) background knowledge. The ubiquitous role of background knowledge lies in the facilitation of meaning making process as it has been shown that learners who have Western background knowledge learn English faster compared to non-Westerners. The processing strategies in the model resemble the generic components of comprehension which can be listed as "grapheme-morphophoneme correspondences, syllable-morpheme information, syntactic information (deep and surface), lexical meaning and contextual meaning, cognitive strategies", and lastly "affective mobilizers" (p. 7). In Kintsch's (1998) CI model of text comprehension, the role of prior knowledge is also predominant as it is expected to be activated during reading when textual propositions interact with the readers' associative networks. In his critical paper on schemata theory in reading comprehension, Nassaji (2007) considers text comprehension as a rather reader-driven process of the integration of textual information into the reader's prior knowledge.

Studies have repeatedly revealed that the amount of recall and comprehension is increased by activating or providing necessary content related knowledge compared to comprehension when no background knowledge is activated nor provided before reading even with two parallel texts (Carrell & Wise, 1998; Alderson, 2000). In an attempt to look for the effects of reading familiar texts, Pulido (2000, 2003) noticed a considerable immediate influence of reading familiar narratives on vocabulary gains as opposed to reading unfamiliar narratives. In a follow-up study, she (2004) investigated the relative effects of reading culturally familiar versus unfamiliar texts on the recognition of nonsense target words by intermediate Spanish learners and elicited greater recognition of words in the culturally familiar text. Although the focus was on the correlation between topic familiarity, text comprehension and form recognition, Leiser's (2007) findings also suggest a significant role of topic familiarity on all dependent variables.

It is apparent from the above review of research that in most studies (e.g., Jacobs et al., 1994; Horst et al., 1998; Rott, 1999) only bivariate correlations were analyzed leaving no room for other factors influential in reading process and vocabulary development, such as background knowledge which can overestimate the role of text comprehension on lexical development. Besides, the studies conducted so far have not made it explicit what kinds and amounts of reading facilitate incidental vocabulary learning. Obviously, the area of incidental vocabulary learning through reading is in need of substantial support from further studies into the extent and nature of word acquisition through reading. As already noted, extensive reading does not always guarantee incidental learning of all frequently exposed unknown words. Building on the common findings stemming from the reviewed literature; that is, the undeniable role of text comprehension in incidental vocabulary acquisition as well as the role of background knowledge in text processing, in comprehension and in short term gains of incidental vocabulary, the present study targets to provide a richer understanding of the nature of these correlations by adopting an experimental design.

### *1.2 Research questions*

- 1) What are the effects of topic familiarity on text comprehension, intake, gain and retention of new vocabulary encountered in narrative passages?
- 2) How and to what extent are text comprehension, lexical intake, gain and retention related?

## **2. Method**

Forty upper-intermediate level learners of English as a foreign language participated in this replication study (i.e., Pulido, 2007). They were freshman students of the English Language Teaching department of a state university in Turkey. All of them were native speakers of Turkish, and their age range was 18-20.

### *2.1 Passages*

Four contrived script-based narrative passages which were adopted from Pulido (2003) are used in the present study. Two of the passages serve as texts on more familiar topics: "The trip to the supermarket" and "The doctor's appointment" whereas the remaining two function as less familiar scenarios: "Publishing an article" and "Buying a house". All of the four stories are in the same length and display a scenario around which the main

character acts making it easier to retain their propositional structures in participants' long-term memory. They also contain similar sentence and discourse level structures, which make them easily comparable

## 2.2 Target Words

From each story, eight words that stand for concepts that are generally associated with narrative scenarios were chosen as target words (henceforth TW) following the operationalization of Pulido (2007) which makes in total thirty-two words. Later, these target words were replaced with nonsense words to serve as targets in order to make sure that no participants have encountered these words before. The nonsense words conformed to the orthographic and morphological rules of English, and the derivational and inflectional suffixes were maintained in inventing them. Additionally, the number of nouns and words was kept constant across stories. Each of the TWs appeared only once in the stories, and contextual clues for guessing their meanings were absent.

## 2.3 Independent Variables

### 2.3.1 Topic Familiarity

Before conducting the study, participants were asked to complete a ten-item Likert scale background questionnaire tapping their degree of familiarity with the given subjects (e.g., playing soccer, building a financial portfolio, etc.). They rated from 1 to 5 their degree of familiarity regarding each topic. The background questionnaire was conducted so as to assure that the categorization of topics as more or less familiar was applicable to this specific participant group. The results of this questionnaire are given in Table 1 below.

Table 1. Descriptive statistics of the background questionnaire

Ratings	More Familiar		Less Familiar					
	Supermarket		Doctor		Article		House	
	<i>f</i>	%	<i>F</i>	%	<i>F</i>	%	<i>F</i>	%
1	1	2.3	6	14.0	15	34.9	16	37.2
2	2	4.7	7	16.3	10	23.3	4	9.3
3	13	30.2	8	18.6	9	20.9	7	16.3
4	12	27.9	14	32.6	5	11.6	10	23.3
5	15	34.9	8	18.6	4	9.3	6	14

Note: Ratings were made on a scale from 1 (not at all familiar) to 5 (extremely familiar).

*N*= 43

The display of the frequencies of responses given to each subject in Table 1 illustrates that the participants are in fact more familiar with the “supermarket” and “doctor” topics, whereas they are less familiar with the “article publishing” and “buying a house” topics.

### 2.3.2 Text Comprehension

The participants were asked to write down in their L1 (Turkish) as much as they could recall from each story immediately after reading them. Propositional segmentation adopted from Pulido (2007) was employed in the scoring of the free recall task. Namely, each story was first divided into propositional units consisting of predicate-argument structures. Maximum points were given for mentioning the gist of the propositions; half points were given for mentioning a fraction of the propositions and no points were given if there are no propositional units. In the end, the total number of recalled propositions by each participant was calculated to reach a comparison of text comprehension of more versus less familiar passages.

### 2.3.3 Time

Incidental vocabulary acquisition of participants was measured on two different occasions: 2 days, and 18 days after reading the passages.

## 2.4 Dependent Variables

### 2.4.1. Intake

To be able to identify lexical items is seen as the first step in the acquisition process, hence in this study, the recognition of TW forms was initially evaluated for revealing intake. Following Pulido (2007), intake was defined as:

a measure of accuracy in memory discrimination for recently processed information—in this case, the TW forms. This task of verifying whether or not certain words were presented within the stories that were read is assumed to tap episodic memory and to be an indirect measure of having noticed the TWs while reading, which should serve as a rough estimate of how deeply textual information was processed (p. 168).

Two days after reading the passages, the participants were asked to indicate on a test of 24 items whether they had seen these words in the passages they read or not. In addition to the eight TWs, from each story eight thematically relevant words are chosen. Lastly, extra eight words related to the theme of each story were given

as distractors in each word recognition text. In total, then, there emerged twenty-four words for which the participants would choose Yes or No options. In order to prevent memory trace effects, those 24 words were randomly distributed within each test. For the assessment of recognition memory test, the participants' choosing Yes options for the TWs was accepted as the hit point.

#### 2.4.2 Lexical Gain and Retention

Incidental vocabulary gain of the participants was assessed via two additional measures two days after reading the passages. The same two tests were again completed by the same participants eighteen days after reading to assess the retention of the TWs. Accordingly, there were eight different tests pertaining to each of the four stories including the eight TWs from each text (since there were two different vocabulary measures, namely, translation production and translation recognition). The order of presentation of lexical items was randomly distributed in order to erase episodic memory traces.

##### 2.4.2.1 Translation Production (L2-L1)

The participants were asked to provide L1 equivalents of TWs of each text both 2 days and 18 days after reading the passages. It is also taken as a measure of productive word knowledge following Pulido (2007).

##### 2.4.2.2 Translation Recognition (L2-L1)

As another measure of receptive vocabulary knowledge, the participants were given a multiple choice test for each narrative passage. This test was also conducted twice. Each test included eight TWs and four L1 equivalents as the options, and a last (Bilmiyorum- Don't Know) option. The options were designed to tap semantic knowledge of the particular words, not the syntactic properties. Within the options, one was the answer, and there were three distractors. The distractors were created in accordance with Pulido's (2007) scheme, that is, a) contextually close word to the TW; b) schematically appropriate word; c) orthographically or phonologically proximate word to another known word in either L1 or L2.

### 2.5 Data Collection and Analysis Procedures

Data were collected in four separate sessions:

Session 1: Topic familiarity (background) questionnaire was administered.

Session 2: One week after the first session, the participants read all 4 stories and completed the written recall test in L1 immediately after each reading.

Session 3: Two days after session 2, one lexical intake and two lexical gain measures were administered.

Session 4: Eighteen days after reading passages, the same two vocabulary measures were administered for assessing lexical retention this time

It is important to note that before reading each passage, the participants were reminded about the comprehension test but the vocabulary tests were unannounced. The design of the present study adopts a repeated measures approach as all participants took part in all experimental conditions; i.e. reading less and more familiar passages. With the help of the propositional analysis of the recall task, all data underwent a quantitative analysis process. Descriptive statistics, Pearson product-moment correlations and *t*-tests were employed to analyze data in the present study.

### 3. Results

In the present study, topic familiarity (more or less familiar narrative passages) and time (2 days and 18 days after reading the stories) act as the categorical independent variables. The purpose of the study was to find out the impact of topic familiarity; that's why it is accepted as the random variable. Reading comprehension, lexical intake (as measured by word recognition test), lexical gain (as measured by translation production and recognition tests), and lexical retention (as measured by the same translation production and recognition tests eighteen days later) served as the dependent variables. Table 2 depicts the descriptive statistics for comprehension scores of more and less familiar passages.

Table 2. Raw means, standard deviations, and range of passage comprehension

Measure	<i>M</i>	<i>SD</i>	Min	Max
More familiar passages	36.05	14.7	16	86
Less familiar passages	21.10	10.15	8	44
Total	28.58	14.64	8	86

*Note.* Values reflect the percentage of semantic propositions correctly recalled. Total = the average of all more and less familiar passages. More familiar = the average of "The Trip to the Supermarket" and "The Doctor's Appointment" passages. Less familiar = the average of "Publishing an Article" and "Buying a House" passages. *N*=40.

In total, it shows that the participants recalled less than half of all possible propositions in their L1 recall protocols. As the overall standard deviation number makes it clear, there is a wide range among the participants in their recalls. Moreover, they remember more semantic propositions from more familiar stories than the less

familiar ones. Also, the statistical analysis of this difference turned out to be significant,  $t(39) = 15.475, p < .001$ . Table 3 shows the descriptive statistics for the dependent variable lexical intake which was measured by the recognition of TW forms the participants encountered two days before.

Table 3. Raw means, standard deviations, and range of lexical intake

Measure	<i>M</i>	<i>SD</i>	Min	Max
More familiar passages	21.82	6.86	4	32
Less familiar passages	21.26	6.79	3	32
Total	21.54	6.79	3	32

*Note.* Values reflect the percentage of semantic propositions correctly recalled. Total = the average of all more and less familiar passages. More familiar = the average of “The Trip to the Supermarket” and “The Doctor’s Appointment” passages. Less familiar = the average of “Publishing an Article” and “Buying a House” passages.  $N=40$ .

As the table makes it obvious, lexical intake from more familiar passages was only slightly higher than less familiar passages. This difference is not statistically significant, though ( $p > .05$ ). The interrelations between passage comprehension and lexical intake were analyzed through the Pearson correlations (see Table 4) to depict whether they are related to each other.

Table 4. Intercorrelations between passage comprehension and lexical intake

Measure	1	2	3	4	5	6
1.Total passage recall	-	.91**	.80**	-.02	.01	.74
2.More familiar passage recall		-	.48**	.04	.05	.02
3.Less familiar passage recall			-	-.11	-.05	-.14
4. Total vocabulary intake				-	.86**	.85**
5.More familiar intake					-	.47**
6.Less familiar intake						-

*Note.* Total = Correlations of all more and less familiar stories. More familiar = “The Trip to the Supermarket” and “The Doctor’s Appointment.” Less Familiar = “Publishing an Article” and “Buying a House.” Recall = Comprehension. \*\* $p < .01$ .  $N=40$ .

Table 4 makes it clear that there are consistent positive correlations within passage recall and lexical intake scores. Accordingly, recall on familiar passages is related with recall on less familiar passages. Likewise, intake from familiar passages correlates with intake from less familiar passages which implies an overall effect of topic familiarity on both text recall and lexical intake. Table 5 presents descriptive statistics for vocabulary gain and retention as measured by translation recognition tests.

Table 5. Raw means, standard deviations, and range of vocabulary gain and retention (translation recognition)

Measure	<i>M</i>	<i>SD</i>	Min	Max
More familiar gain	8.90	2.07	4	13
Less familiar gain	5.27	2.10	1	9
Total vocabulary gain	7.13	2.76	1	13
More familiar retention	7.97	2.77	2	12
Less familiar retention	6.30	2.08	2	11
Total vocabulary retention	7.15	2.58	2	12

*Note.* Vocabulary gain and retention maximum per story = 8. Gain = 2 days after reading. Retention = 18 days after reading. Total = the average of all more and less familiar stories. More familiar = “The Trip to the Supermarket” and “The Doctor’s Appointment.” Less familiar = “Publishing an Article” and “Buying a House.”  $N=40$ .

The table shows that 2 days after reading participants recognized more TW translations from more familiar passages than less familiar ones. Paired samples t-tests also proved that there was a significant difference between more and less familiar vocabulary gains,  $t(39) = 7.143, p < .001$ . However, in the second condition, 18 days after reading, this difference is not significant ( $p > .05$ ).

Table 6. Intercorrelations among passage comprehension and, vocabulary gain and retention (translation recognition)

Measure	1	2	3	4	5	6	7	8	9
1.Total passage recall	-	.91**	.80**	-.37**	-.27	-.40*	-.23	.02	-.33
2.More familiar passage recall		-	.48**	-.43**	-.33*	-.42*	-.18	.05	-.31
3.Less familiar passage recall			-	-.14	-.09	-.21	-.18	-.04	-.20
4.Total gain				-	.82**	.83**	.31	-.01	.40*
5.More familiar gain					-	.37*	.14	-.19	.32
6.Less familiar gain						-	.36	.17	.34
7.Total retention							-	.66**	.70**
8. More familiar retention								-	-.05
9. Less familiar retention									-

Note. Total = Collapse of all more and less familiar stories. More familiar = “The Trip to the Supermarket” and “The Doctor’s Appointment.” Less Familiar = “Publishing an Article” and “Buying a House.” Recall = Comprehension. \* $p < .05$ . \*\* $p < .01$ .  $N=40$ .

Table 6 demonstrates how passage comprehension, vocabulary gain and vocabulary retention are related as measured through translation recognition tasks. Overall, there is a strong positive correlation between passage comprehension and topic familiarity. Moreover, the bivariate correlations illustrated that there is a modest negative correlation between more familiar passage comprehension and more familiar vocabulary gain. That is, the participants recognized fewer translation equivalents in more familiar passages although their recall was high. Furthermore, there is a moderate negative correlation between more familiar recall and less familiar vocabulary gain suggesting that the students’ recognition of target words of the less familiar passages decreased as their recall of more familiar passages increased. Besides, there is a statistically significant negative correlation between total recall and total vocabulary gains which implies that the participants did not gain much from passage comprehension. For vocabulary retention as measured by translation recognition, no significant correlations were observed except for a correlation in itself. Yet, there exists a moderate positive correlation between total lexical gain and less familiar vocabulary retention which can imply that the participants’ recall of more TW translations in total on the second day led to an increased vocabulary retention on less familiar condition on the eighteenth day.

Table 7 displays the descriptive statistics for vocabulary gain and retention scores as measured by translation production tasks.

Table 7. Raw means, standard deviations, and range of vocabulary gain and retention (translation production)

Measure	<i>M</i>	<i>SD</i>	Min	Max
More familiar gain	5.97	3.03	11	13
Less familiar gain	3.37	2.19	0	8
Total vocabulary gain	4.71	2.94	0	13
More familiar retention	7.64	2.49	2	12
Less familiar retention	3.89	2.29	1	9
Total vocabulary retention	5.92	3.040	1	12

Note. Vocabulary gain and retention maximum per story = 8. Gain = 2 days after reading. Retention = 18 days after reading. Total = the average of all more and less familiar stories. More familiar = “The Trip to the Supermarket” and “The Doctor’s Appointment.” Less familiar = “Publishing an Article” and “Buying a House.”  $N=40$ .

On average, the participants produced more TW translations on both Time 1 and Time 2 in familiar passage conditions. For both measures, there are significant differences between the means of more familiar vs. less familiar texts,  $t_{\text{gain}}(39) = 6.261, p < .05$ ;  $t_{\text{retention}}(39) = 8.582, p < .001$ . Interestingly, there was a slight increase in the total mean of retention scores compared to the total mean of gain scores. These differences between immediate and delayed testing times were not statistically significant, though ( $p > .05$ ). Pearson’s product-moment correlations were computed to understand the relationships among text comprehension and productive vocabulary gain and retention (See Table 8).

Table 8. Intercorrelations among passage comprehension and, vocabulary gain and retention (translation production)

Measure	1	2	3	4	5	6	7	8	9
1.Total passage recall	-	.91**	.80**	.06	.12	-.16	-.23	.02	-.33
2.More familiar passage recall		-	.48**	.07	.14	-.16	-.18	.05	-.31
3.Less familiar passage recall			-	.02	.03	-.09	-.18	-.04	-.20
4.Total gain				-	.84**	.65**	-.13	-.02	-.13
5.More familiar gain					-	.13	-.26	-.14	-.20
6.Less familiar gain						-	.13	.13	.04
7.Total retention							-	.66**	.70**
8. More familiar retention								-	-.05
9. Less familiar retention									-

*Note.* Total = Collapse of all more and less familiar stories. More familiar = “The Trip to the Supermarket” and “The Doctor’s Appointment.” Less Familiar = “Publishing an Article” and “Buying a House.” Recall = Comprehension. \*\* $p < .01$ .  $N=40$ .

In Table 8, the moderate positive correlation between passage recall from more familiar text and passage recall from less familiar text suggests an effect of topic familiarity. Apart from the consistent correlations within the discrete measures (i.e., passage recall, vocabulary gain, and vocabulary retention), no significant relationships emerged between the variables.

#### 4. Discussion

This study was an attempt to investigate whether there is a relationship between text comprehension and incidental vocabulary gains. It also seeks out if there is, how topic familiarity influences the reciprocal relationship between passage comprehension and lexical gain. Overall, the results of this research presented that there is not a significant relationship between comprehension and vocabulary gains. To the contrary, mostly a negative relationship was observed. For instance, there exists a negative relationship between passage comprehension of more familiar texts and total vocabulary gains. This shows that the participants of this study could not recognize the translations of nonsense TWs even if they understood and recalled the stories. Moreover, there was not detected any significant relationship between more familiar passage recall and vocabulary retention in more familiar passages, again proposing that the participants’ comprehension of the stories did not affect their vocabulary recognition. The same pattern emerged repeatedly across the tasks. In sum, this suggests that lexical gain acted independently from passage comprehension in this study. Namely, the answer to the first research question is that there is not a significant relationship between text comprehension and vocabulary gain.

In this regard, this study does not confirm the conclusions of the study it was replicating. Even though Pulido (2007) found a consistently significant effect of passage comprehension on all levels of incidental vocabulary acquisition, they were not related to each other on any measure in the present study. This may partially stem from the higher proficiency levels of the participants in the present. Unlike Pulido’s (2007), the participants of this study did not have much difficulty in understanding the stories and writing the summaries of any of the texts despite the presence of nonsense words. As a result, although their performance on text recall changed from the more familiar to less familiar conditions, their vocabulary gains did not change according to their comprehension of the texts.

On the other hand, topic familiarity certainly functioned as a reliable predictor for all of the measures in the present study. The participants gained significantly more sight vocabulary from the more familiar passages compared to the less familiar ones. They also recalled significantly more propositions from more familiar texts. The only condition where topic familiarity did not play a significant role was lexical intake measures. The participants identified almost the same number of TWs from less familiar passages as more familiar ones. It somehow confirms Pulido’s (2007) findings since in her study also the participants could not adeptly discriminate TWs from nontargets in more familiar conditions. She interpreted this as an impact of interference or source confusion because it was shown that when participants read script-based narratives that are about familiar topics, there might be less vocabulary intake. There are two sources of declarative memory which are episodic memory and semantic memory. As the participants were asked to recognize TWs two days after reading, they were obviously retrieving information from their declarative memory sources. Hence, semantic memory sources might have interfered with episodic memory sources and led to the incorrect detection of words. Furthermore, as Pulido (2007) maintains the distractors in the present story were also thematically related words, so it was more than natural to expect source confusion in the case of familiar scenarios. For less familiar script-based narrative passages, this would not create a huge problem as they already possess less semantic knowledge



regarding the subject.

Apart from these, topic familiarity positively correlated with the other two lexical gain measures; i.e., translation recognition and production. In line with Pulido's (2007) findings, the participants in this study recognized and produced more translation equivalents of TWs in more familiar topics as opposed to less familiar narratives. Similarly, a moderately significant relationship between passage comprehension and topic familiarity was observed. Obviously, these findings also lend support to De Bot, Paribakht, and Wesche's (1997) model of lexical inference which proposes lemma component (i.e., necessary background knowledge) as central to mental lexicon. They put forward that learners process unknown words more easily if they could find the related networks in their mental lexicon. It also provides further support to Mondria and Wit-de Boer's (1991) study, which contended that contextually elaborate ("pregnant" in their words) texts are more conducive to correct guessing the meaning of unknown words.

## 5. Conclusions

This study proved it once again that topic-related background knowledge facilitates both comprehension and incidental vocabulary development (Coady, 1993; De Bot, Paribakht, & Wesche, 1997; Meara, 1997; Mondria & Wit-de Boer, 1991). Hence, to possess a text related prior knowledge seems as one of the prerequisites of reading classes for consolidating the reciprocal relationship of topic familiarity and lexical development through comprehension. It is advisable that teachers should provide or activate the learners' scripts regarding the topic when necessary so that it can facilitate both incidental vocabulary learning and comprehension process.

There are certain limitations of the present study that can be taken into consideration while drawing conclusions from the results. To start, this study was conducted only with higher proficiency level learners, as a result there did not appear a strong effect of passage comprehension on vocabulary development measures since they might have already been adept in comprehension skills. Furthermore, the data collection procedure of the present study overwhelmed the participants, and this was most probably reflected in the results. Asking the participants to write as much as they recalled from all of the four stories at once might have distressed and bored them and, in turn, affected their performance. In the future, the participants may be requested to read the passages at different but close time slots in order to prevent the effects of burnout.

## References

- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Anderson, N. J. (1999). *Exploring second language reading: Issues and strategies*. Boston, MA: Heinle and Heinle.
- Bernhardt, E.B. (1984) Toward an information processing perspective in foreign language reading. *The Modern Language Journal*, 68, 323-331.
- Bernhardt, E. B. (1991). *Reading development in a second language: Theoretical, empirical, and classroom perspectives*. Norwood, NJ : Ablex.
- Bernhardt, E. B. & Kamil, M. L. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypotheses. *Applied Linguistics*, 16, 15-34.
- Carrell, P., & J. C. Eisterhold (1988). Schema theory and ESL reading pedagogy. In P. L. Carrell, J. Devine, and D. E. Eskey (Eds), *Interactive approaches to second language reading* (pp. 73-92). Cambridge: Cambridge University Press.
- Carrell, P., & Wise, T. (1998). The relationship between prior knowledge and topic interest in second language reading. *Studies in Second Language Acquisition*, 20, 285-309.
- Coady, J. (1979). A psycholinguistic model of the ESL reader. In R. Mackay, B. Barkman, and R.R. Jordan (Eds.), *Reading in a second language* (pp. 5-12). Rowley, Mass: Newbury House.
- Coady, J. (1993). Research on ESL/EFL vocabulary. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 3-23). Norwood, NJ: Ablex.
- Davis, F.B. (1968) Research in comprehension in reading, *Reading Research Quarterly*, 3, 499-545.
- De Bot, K., Paribakht, T., & Wesche, M. (1997). Toward a lexical processing mode for the study of second language vocabulary acquisition. *Studies in Second Language Acquisition*, 19, 309-329.
- Eskey, D. E. (1988). Holding in the bottom: an interactive approach to the language problems of second language readers. In P. L. Carrell, J. Devine, and D. E. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 93-100). Cambridge: Cambridge University Press.
- Goodman, K. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist* 6, 126-35.
- Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly*, 25 (3), 375-405.
- Grabe, W. (2004). Research on teaching reading. *Annual Review of Applied Linguistics*, 24, 44-69.
- Grabe, W. & Stoller, F. (1997). Reading and vocabulary development in a second language. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 98-122). New York: Cambridge University Press.

- 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.
- Hudson, T. (2007). *Teaching second language reading*. New York: Oxford.
- Jacobs, G., Dufon, P., & Fong, C. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading, 17*, 19–28.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Kintsch, W. (2004). The construction-integration model of text comprehension and its implications for instruction. In R. B. Ruddell and N. J. Unrau (Eds.), *Theoretical model and processes of reading, 5th edition* (pp. 1270-1329). Newark, DE: International Reading Association.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363–394.
- Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge: Cambridge University Press.
- Koda, K. (2007). Reading and language learning: Crosslinguistic constraints on second language reading development. *Language Learning, 57(1)*, 1-44.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *Modern Language Journal, 73*, 440–464.
- Leeser, M. J. (2007). Learner based factors in L2 reading comprehension and processing grammatical form: Topic familiarity and working memory. *Language Learning, 57(2)*, 229-270.
- Maria, K. (1990). *Reading comprehension instruction: Issues and strategies*. Michigan: York Press.
- Meara, P. (1997). Towards a new approach to modeling vocabulary acquisition. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 109–139). Cambridge: Cambridge University Press.
- Mondria, J., & Wit-de Boer, M. (1991). The effects of contextual richness on the guessability and retention of words in a foreign language. *Applied Linguistics, 12*, 249–267.
- Nagy, W. E., Herman, P. A., & Anderson, R.C. (1985). Learning words from context. *Reading Research Quarterly, 20(2)*, 233-253.
- Nassaji, H. (2007). Schemata theory and knowledge-based processes in second language reading comprehension: A need for alternative perspectives. In K. Koda (Ed.), *Reading and language learning* (pp. 79-113). Oxford: Blackwell Publishing.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nuttall, C. (2005). *Teaching reading skills in a foreign language*. Oxford: Macmillan Education.
- Paribakht, T., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174-200). New York: Cambridge University Press.
- Pulido, D. (2000). *The impact of topic familiarity, L2 reading proficiency, and L2 passage sight vocabulary on incidental vocabulary gain through reading for adult learners of Spanish as a foreign language*. Unpublished doctoral dissertation. University of Illinois at Urbana-Champaign.
- Pulido, D. (2003). Modeling the role of second language proficiency and topic familiarity in second language incidental vocabulary acquisition through reading. *Language Learning, 53*, 233–284.
- Pulido, D. (2004). The effect of cultural familiarity on incidental vocabulary acquisition through reading. *The Reading Matrix: An Online International Journal, 4*, 20–53.
- Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies in Second Language Acquisition, 21*, 589–619.
- Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.) *Attention and Performance*. VI. Hillsdale, N. J.: Lawrence Erlbaum Associates.
- Rumelhart, D.E. (2004). Toward an interactive model of reading. In R.B. Ruddell, & N.J. Unrau (Eds.), *Theoretical Models and Processes of Reading* (pp. 1149-1179). Newark, DE: International Reading Association.
- Salmerón, L., Kintsch, W. & Cañas, J. J. (2006). Reading strategies and prior knowledge in learning from hypertext. *Memory and Cognition, 34(5)*, 1157-1171.
- Sinatra, G. M., K. J. Brown & R. E. Reynolds (2002). Implications of cognitive resource allocation for comprehension strategies instruction. In C. C. Block and M. Pressley (Eds.), *Comprehension instruction* (pp. 62-76). New York: The Guilford Press.
- St. John, M. F., & McClelland, J. L. (1992). Parallel constraint satisfaction as a comprehension mechanism. In R. G. Reilly & N. E. Sharkey (Eds.), *Connectionist approaches to natural language processing* (pp. 97–136). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Stanovich, K.E. (1980) Toward an interactive-compensatory model of individual differences in the development

of reading fluency, *Reading Research Quarterly*, 16, 32-71.

Urquhart, S., & C. Weir (1998). *Reading in a second language: Process, product and practice*. London: Longman.

#### **Notes**

Note 1. This paper was presented at the 2<sup>nd</sup> International Symposium on Language Education and Teaching in Rome, Italy on 20-23 April, 2017.