Original Research

Vocabulary and Reading Comprehension Revisited: Evidence for High-, Mid-, and Low-Frequency Vocabulary Knowledge

Ahmed Masrai

Abstract
The association between vocabulary knowledge and reading comprehension has been extensively researched. However, modeling the contribution of vocabulary knowledge within different frequency ranges to second language (L2) learners’ reading comprehension is an underexplored area. Thus, the present study examines the degree to which high-, mid-, and low-frequency-based levels of orthographic vocabulary knowledge are able to predict L2 reading comprehension. A vocabulary size test and the reading section of International English Language Testing System (IELTS) were administered to 256 tertiary-level Arab learners of English. The participants’ language proficiency ranged from B2 to C1 of Common European Framework of Reference (CEFR) levels. Results showed that high- and mid-frequency word ranges contributed uniquely to the L2 reading comprehension for the entire cohort. When the participants were categorized to relatively low- and relatively high-proficiency subgroups, only high-frequency range explained variance in L2 reading comprehension for the low-proficiency subgroup. Among the high-proficiency subgroup, high-, mid-, and low-frequency-based ranges offered unique contribution to L2 reading comprehension, but mid-frequency range explained the largest variance. The findings provide evidence aimed at informing approaches to the development of overall vocabulary size and the mid-frequency words, and not just a focus on the most frequent vocabulary, for the purpose of supporting L2 reading comprehension.

Keywords
L2 reading comprehension, orthographic vocabulary knowledge, vocabulary size, frequency levels, vocabulary and reading

Introduction
Research on the relationship between vocabulary size and language proficiency in second language (L2) learners has been extensively conducted within the realm of reading. Many studies have shown strong positive correlations between receptive vocabulary knowledge and reading comprehension, ranging from .50 to .85, among L2 learners of different proficiency levels (e.g., Henriksen, Albrechtsen, & Haastrup, 2004; Laufer, 1992a; Milton, Wade, & Hopkins, 2010; Qian, 2002). This robust association between receptive vocabulary knowledge and reading comprehension has led many researchers to accentuate that vocabulary size is the determinant factor for reading achievement in the L2 context. However, for L2 language teachers and learners, in addition to knowing the volume of vocabulary needed to read authentic materials, an important question is, “What type of vocabulary (i.e., high-, mid-, or low- frequency words) is more instrumental in comprehending a written text?”

Although a substantial number of studies have found vocabulary knowledge to be a significant predictor of reading success in L2 learners and have established certain vocabulary size and lexical coverage targets for comprehension (e.g., Hazenberg & Hulstijn, 1996; Hu & Nation, 2000; Laufer, 1992a; Nation, 2006; Schmitt, Jiang, & Grabe, 2011), most of those studies have predominantly focused on the relationship between the overall vocabulary size and reading comprehension. There are key areas of research which have yet to be adequately examined. To develop evidence-based frameworks for L2 vocabulary for reading programs, a better understanding of the interaction between L2 orthographic vocabulary knowledge (OVK) and L2 reading comprehension within cohorts of L2 learners is needed. The present study, therefore, is an attempt to empirically disentangle the contribution of three levels of word frequency namely, high frequency, mid frequency, and low frequency to reading comprehension in L2 learners. The study is particularly

1King Abdul Aziz Military Academy, Riyadh, Saudi Arabia

Corresponding Author:
Ahmed Masrai, Department of Languages, King Abdul Aziz Military Academy, Riyadh 11583, Saudi Arabia.
Email: a.masrai@hotmail.com
motivated by a scarcity of empirical research on the influence of vocabulary knowledge of specific word frequency levels on L2 reading comprehension. The present study seeks to begin filling this gap in reading literature.

**Literature Review**

**Vocabulary Knowledge and Reading Comprehension**

Reading comprehension encompasses abilities to recognize words promptly and efficiently, develop and use a wide range of recognition vocabulary, process sentences to build comprehension, engage a variety of strategic processes and underlying cognitive skills, interpret and evaluate texts matching reader targets and needs, and process texts fluently over a protracted period of time (Grabe, 2009). These kinds of processes and knowledge resources allow the reader to effectively generate written discourse comprehension to the desired level. Among the many variables involved in comprehending a written text is vocabulary knowledge. The strong association between a learner’s vocabulary knowledge and the ability to effectively negotiate L2 reading tasks emphasizes the importance of L2 learners having adequate levels of vocabulary knowledge to allow coping with the linguistic demands of this essential L2 skill (e.g., Nation, 2001; Qian, 2002; Stæhr, 2008, 2009).

Several studies have provided valuable insights into the link between L2 vocabulary knowledge and L2 reading comprehension. For example, Laufer (1992a, 1992b) investigated the relationship between receptive vocabulary knowledge and reading comprehension and found a relatively high correlation, ranging from .50 to .75 between the two factors. Qian (2002), in the same vein, examined the relationship between vocabulary knowledge and reading among 217 L2 learners of English with a wide range of first language (L1) backgrounds. Findings of Qian’s study indicated that a strong and significant correlation existed between vocabulary size and reading comprehension ($r = .74, p < .01$). Li and Kirby (2015) also explored the relation between vocabulary knowledge and reading comprehension among 246 Chinese school learners of English using different measures of vocabulary knowledge, including breadth and depth vocabulary knowledge. Their findings showed that both the breadth and depth dimensions of vocabulary knowledge correlated significantly with the scores for reading.

Few studies have examined the relationship between vocabulary knowledge and the skill of reading among a single cohort of L2 learners (Cheng & Matthews, 2018). Stæhr (2008), for example, examined the connection between L2 vocabulary knowledge of 88 Danish learners of English as a foreign language (EFL) and their L2 reading comprehension. Receptive OVK and reading comprehension were found to be strongly correlated ($r = .83, p < .01$). Stæhr’s study, interestingly, revealed that vocabulary size explained about 72% of the variance in reading comprehension scores. Another contribution toward understanding the relationship between vocabulary knowledge and reading comprehension was made by Milton et al. (2010). The authors used scores on an orthographic vocabulary measure (X-Lex; Meara & Milton, 2003) and reading scores on the International English Language Testing System (IELTS) of learners from different L1 backgrounds, including Arabic, Chinese, Japanese and others from European countries, to quantify the link between vocabulary size and reading comprehension. The results of the study indicated that vocabulary size correlated strongly with reading ($r = .70, p < .01$), explaining about 50% of the variance.

In sum, the literature shows some levels of consistent reliable correlations between vocabulary knowledge and reading comprehension, and that vocabulary is a reliable predictor of reading ability. Thus, readability indexes, not surprisingly, include vocabulary as a major element, indicating that word knowledge affects text comprehension to a large extent (e.g., Graves, 1986; Stahl, 2003). To further explore the relationship between vocabulary and reading, the present study seeks to broaden our understanding about the contribution of three frequency levels (high, mid, and low) of vocabulary to reading comprehension. Different from most of the previous studies, which have looked at the relation between the overall vocabulary size and reading, the current study treats knowledge of high-, mid-, and low-frequency levels as three unique independent variables on reading comprehension.

**Lexical Coverage and Reading Comprehension**

Studies that have been devoted to investigating the connection between vocabulary and reading comprehension have introduced the notion of ‘lexical threshold’—that is, the volume of vocabulary required before higher-level comprehension strategies can be operationalized (e.g., Hu & Nation, 2000; Laufer, 1989, 1992a; Nation, 2006; Schmitt et al., 2011). The notion of lexical threshold centers around two related factors: (a) lexical coverage—the percentage of known words learners understand in a given text and (b) receptive vocabulary knowledge they need to attain this coverage. For example, if a learner’s lexical coverage of a given text is 90%, this means that her or his understanding of the running words in the text is 90%.

Earlier studies (e.g., Laufer, 1989, 1992a) suggested that around 3,000-word families can provide the lexical coverage that is required to read authentic materials independently. However, in a later study, Hu and Nation (2000) reported that participants in their study needed to know 98% to 99% coverage of a written text before adequate comprehension was possible. Currently, the consensus appears to be that an optimal coverage for reading of any text is 98% of word tokens and the minimal coverage is 95% (Laufer & Ravenhorst-Kalovski, 2010). In this vein, a number of researchers have established vocabulary size figures related to different text.
First, high-frequency vocabulary has been defined to include parameters of high-, mid-, and low-frequency vocabulary. The purpose of this section was to determine the most useful high-, mid-, and low-frequency vocabulary in reading is explained by different levels of vocabulary type. However, the contribution of words at different frequency levels to a reading comprehension model is yet to be examined.

Second, according to Schmitt and Schmitt (2014), mid-frequency vocabulary includes words beyond the 3,000 level and less than the 10,000 level, that is 3,001 to 9,000. The best way to describe mid-frequency vocabulary is by citing examples from Schmitt and Schmitt (2014), who were the first to clearly introduce this label. The examples explain how mid-frequency vocabulary relates to language use. Table 1 exemplifies the type of vocabulary at each 1,000 level in the mid-frequency band.

Schmitt and Schmitt (2014) pointed out that it is definitely worth learning mid-frequency words such as these presented in Table 1, because research has demonstrated that accumulating increasing amounts of vocabulary in the mid-frequency range leads to very clear rewards. One very essential reward is getting engaged with English for authentic purposes, such as watching movies. In two studies by Webb and Rodgers (2009a, 2009b), it was determined that knowledge of the most frequent 3,000-word families gives a little over 95% coverage in a range of television programs and movies. Although this coverage may offer a reasonable level of comprehension, there remains about 4% to 5% of unknown words, which account for around 3.9 unknown vocabulary items per minute (Schmitt & Schmitt, 2014, p. 495). As the purpose of watching movies is typically pleasure, the number of unknown words may affect viewing, and thus enjoyment. This would

### Table 1. Example Words From Different 1,000 Level of Mid-Frequency Band.

<table>
<thead>
<tr>
<th>Mid-frequency level</th>
<th>Example words</th>
</tr>
</thead>
<tbody>
<tr>
<td>3,001-4,000</td>
<td>Academic, consist, exploit, rapid, vocabulary</td>
</tr>
<tr>
<td>4,001-5,000</td>
<td>Agricultural, contemporary, dense, insight, particle</td>
</tr>
<tr>
<td>5,001-6,000</td>
<td>Cumulative, default, penguin, rigorous, schoolchildren</td>
</tr>
<tr>
<td>6,001-7,000</td>
<td>Axis, comprehension, peripheral, sinister, taper</td>
</tr>
<tr>
<td>7,001-8,000</td>
<td>Authentic, conversely, latitude, mediation, undergraduate</td>
</tr>
<tr>
<td>8,001-9,000</td>
<td>Anthropology, fruitful, hypothesis, semester, virulent</td>
</tr>
</tbody>
</table>

Source. Adapted from Schmitt and Schmitt (2014, p. 495)
also be the case when reading authentic materials or reading for pleasure, which is another important reward of learning mid-frequency vocabulary. In their studies, Webb and Rodgers (2009a, 2009b) argued that mid-frequency vocabulary was the important range of words that enhanced comprehension of various types of television programs. For reading comprehension, there is almost a consensus that knowledge of the 8,000- to 9,000-word families is required for reading authentic materials (e.g., Nation, 2006; Schmitt et al., 2011). However, precisely what level of contribution does mid-frequency vocabulary provide along the continuum of word frequency range remains to be investigated.

Finally, low-frequency vocabulary includes the words over 9,000-word families. At this level of word frequency, vocabulary becomes very infrequent and thus has very limited use. Evidence for the limited utility of low-frequency vocabulary is found in Nation’s (2006) study. Nation analyzed a range of English authentic materials (i.e., novels, newspapers) and quantified that it requires knowledge of the most frequent 8,000- to 9,000-word families, plus proper nouns, to reach the 98% coverage of written texts. This coverage is thought to facilitate efficient reading. If knowledge of the 8,000- to 9,000-word families is sufficient for reading comprehension of a wide range of texts without being disproportionately constrained by a lack of vocabulary knowledge, then intentional focus on teaching/learning low-frequency vocabulary might not be so important, at least for general language use.

**Vocabulary Measures and Reading Comprehension**

The two notions of vocabulary that are widely reported in the literature in relation to reading comprehension are breadth and depth of vocabulary knowledge. While the definition of breadth knowledge construct is less controversial among researchers, defining depth knowledge is far more complex. Briefly, vocabulary breadth refers to “the number of words for which the person knows at least some of the significant aspects of meaning,” while vocabulary depth elaborates on “how well a learner knows a given word” (Anderson & Freebody, 1981, p. 93). In other words, breadth of vocabulary knowledge refers to how many words a person knows whereas depth of vocabulary knowledge refers to how well a person knows these words. While counting how many words a person knows might appear straightforward, identifying how well a person knows a word is very problematic. This complexity of defining the depth construct has led some researchers (e.g., Li & Kirby, 2015; Tannenbaum, Torgesen, & Wagner, 2006) to use different measures in a single study to tap into the depth knowledge when examined in relation to reading comprehension.

To this end, the present study opted for using a breadth measure of vocabulary for two main reasons. First, because there is less agreement in the literature about the defining construct of depth of vocabulary knowledge and, second, because recognition vocabulary (receptive knowledge) has been well established as a good predictor of reading comprehension (e.g., Laufer & Ravenhorst-Kalovski, 2010; Nation, 2006; Nguyen & Nation, 2011). Laufer and Avidan-Levitzky (2017) argued that due to the impact of vocabulary on reading, determining learners’ vocabulary size is useful when planning reading programs and when assigning learners to the appropriate proficiency level. They also pointed out that vocabulary knowledge needed for reading is receptive rather than productive inasmuch as readers only need to understand the meaning of a word in a given text (Laufer, & Avidan-Levitzky, 2017).

Since receptive vocabulary size is very crucial for reading comprehension, measures such as vocabulary levels test (VLT; Nation, 1983; Schmitt, Schmitt, & Clapham, 2001) and vocabulary size test (VST; Nation & Beglar, 2007) may be considered adequate for testing reading vocabulary because they measure the meaning recognition of words sampled from different frequency levels (Laufer & Avidan-Levitzky, 2017). For the purpose of the current study, the VST was used because it covers a wide range of frequency bands, including high-, mid-, and low-frequency levels, which are the focus of the present study.

As the VST was used in this study, it is discussed in some detail. The VST includes samples of words from 14 sequential frequency bands (first 1,000-fourteenth 1,000), with 10 items representing each band. The total number of items included in the VST is 140, which measures vocabulary knowledge of 14,000-word families. Since the test comprises equally distributed items across the frequency bands, items sampled from each band should give an estimation of word knowledge in that band. The target items of the VST are presented in nondefining minimal context sentences, followed by four definitions for a given item where a learner needs to choose the correct one. The following example shows an item from the VST:

**JUMP:** She tried to <jump>.

1. Lie on top of the water;
2. Get off the ground suddenly;
3. Stop the car at the edge of the road;
4. Move very fast.

The VST has been widely used as a diagnostic and a research tool, particularly after evidence for the construct validity of the test was demonstrated by Beglar (2010) using the Rasch modeling (e.g., Bundgaard-Nielsen, Best, & Tyler, 2011; Coxhead, Nation, & Sim, 2015; Elgort, 2013; Schmitt & Schmitt, 2014). In his study, Beglar (2010) found that the majority of the test items fit the Rasch model and the test as a whole is a unidimensional measure of receptive vocabulary knowledge. After establishing the strong reliability for the VST, Beglar (2010) concluded that “test-takers were measured with a high degree of precision on multiple versions of
the test” (p. 116). In a later study, Leeming (2014) examined the concurrent validity of the VST and found statistically significant correlations between the learners’ scores on the VST and reading comprehension test, offering support for the concurrent validity of the VST when reading was used as a criterion. However, in spite of the reliability and validity evidence for the VST, it has been criticized on two grounds. One is that the VST may overestimate test-takers’ vocabulary size, as it is a multiple-choice test (McLean, Kramer, & Stewart, 2015; Stewart, 2014), and the other is its suitability for measuring reading comprehension (Gyllstad, Vilkaitė, & Schmitt, 2015).

Taking these criticisms into account, Laufer and Aviad-Levitzky (2017) examined the adequacy of the VST as a measure of learners’ vocabulary size and its predictive power for reading ability. They used two measures of vocabulary knowledge in relation to reading comprehension, the VST as a measure of recognition vocabulary and a meaning recall test. Their findings revealed that although both measures predicted reading ability, the VST performed slightly better at times. Laufer and Aviad-Levitzky (2017) also concluded that word recognition tests, such as VST, are “highly suitable as measures of vocabulary size and predictors of reading comprehension” (p. 739).

**Gap in the Literature**

Although vocabulary knowledge has been recognized as a good predictor of reading comprehension, as demonstrated in the above review, the precise nature of the contribution of high-, mid-, and low-frequency vocabulary to reading comprehension is still not clear. To the researcher’s knowledge, there are a few or no studies on the effect of these three variables on reading comprehension in L2 learners. Most L2 studies have examined the relationship between reading comprehension and the overall vocabulary size of learners without disentangling the contribution of high-, mid-, and low-frequency vocabulary to the L2 reading model.

**The Study**

The present study was designed to fill a gap in the literature on L2 reading. It examines the contribution of vocabulary knowledge at three frequency levels (high, mid, and low) to reading comprehension among L2 learners of English. Based on the literature reviewed, the VST is considered a suitable measure of vocabulary knowledge in relation to reading comprehension. To quantify the effect of high-, mid-, and low-frequency vocabulary on reading comprehension, scores on bands contributing to each frequency level were summed together. The three frequency levels were referred to in the study as OVK1, OVK2, and OVK3, representing high-, mid-, and low-frequency words, respectively. For example, OVK1 is the total score of the first three frequency bands of the VST (1,000-3,000); OVK2 is the total score of 4,000 to

9,000; and OVK3 is the total score of 10,000 to 14,000. To this end, two research questions were addressed to achieve the aim of the study: (1) What is the contribution of OVK of high-, mid-, and low-frequency words to the prediction of L2 reading comprehension for a single cohort of learners? (2) What is the contribution of OVK of high-, mid-, and low-frequency words to the prediction of L2 reading comprehension for learners of relatively low- and relatively high-L2 proficiency levels?

**Method**

**Participants**

The participants in the present study were 256 adult students from three higher education institutions in Riyadh province, Saudi Arabia. The students were from the first to fourth year of university study majoring in English language programs. The participants’ English language proficiency ranged from B2 to C1 levels of Common European Framework of Reference (CEFR) at the time of the data collection. The total group consisted of 112 females and 144 males, with ages between 18 and 22. All the participants have Arabic as their first language and had received on average 11 years of English language study. Participation was voluntary, and the study was conducted in accordance with the institutions’ ethical guidelines.

**Measures**

**OVK test.** The VST (Nation & Beglar, 2007) was used to measure the participants’ vocabulary size. This receptive vocabulary measure is widely used in studies of L1 and L2 acquisition and involves vocabulary drawn from different frequency bands in English using the BNC as the basis of word selection. The test has been reported in the literature as a valid and reliable measure of written vocabulary knowledge (see, for example, Beglar, 2010). The 14,000-word version of the VST contains 140 multiple-choice items, with 10 items from each 1,000-word family level. The items are grouped on the test according to their frequency, high to low (i.e., first 1,000-fourteenth 1,000). The number of items that a test-taker chooses correctly is multiplied by 100 to calculate their total receptive vocabulary size. The VST measures learners’ knowledge of words in the orthographic form, that is, the link of form and meaning, and to a smaller degree concept knowledge (Nation, 2012). The test items appear in a single nondefining context in English, which allows learners to gauge the word class to which the word belongs but does not provide further cues to the word meaning. Test-takers are required to choose the corresponding meaning from four given choices. The VST has the merit of covering the high-, mid-, and low-frequency words, which allowed purposeful categorization of the frequency levels examined in the current study.
To this end, in addition to the interest in knowing the total vocabulary score of each participant for the purpose of grouping the participants to either high- or low-proficiency level, the main objective was obtaining the scores corresponding to each frequency band. Therefore, scores of the participants from the first 1,000 to the third 1,000; from the fourth 1,000 to the ninth 1,000; and from the tenth 1,000 to the fourteenth 1,000 were computed separately to form the high-, mid-, and low-frequency levels (referred to as OVK1, OVK2, and OVK3, respectively). Categorization of these three frequency levels was sought to fulfill the study aim of examining the contribution of knowledge at each level to L2 learners’ reading comprehension.

**Reading comprehension test.** A version of the reading sub-section of IELTS (see the Appendix) was used to measure the informants’ L2 reading comprehension, as in Milton et al. (2010). IELTS is a standardized measure of overall English language proficiency and is widely used as a standard measure for admission to many universities around the world. To this end, the participants were required to answer 40 questions of varied types, including multiple choices, identifying information, identifying the writer’s views/claims, matching information, matching headings, sentences completion, summary completion, and table completion. Each correct answer was worth one point. The participants were given 60 min to complete the test, including the time given to transfer answers to the answer sheet provided with the questions’ booklet. The participants were also directed to be careful when writing their answers on the answer sheet as poor spelling would be taken into account when marking the test. It should also be noted here that although the participants in the present study were familiar with the task format (i.e., types of reading comprehension questions), the topics of reading might not be familiar to all of them. This, however, was noted in the “Limitations” section.

**Procedure**

**Tests administration.** The two tests were delivered in two sessions after prior arrangements with the participants and their institutions had been made. The overall data collection lasted approximately 1 hr and 50 min. The participants were given clear information about the purpose of the present study and also clear instructions on how to perform the tasks. During the testing sessions, participants were not allowed to use any electronic devices and were instructed to complete the tests in silence complying with the given time frames.

**Categorization of Low- and High-Proficiency Subgroups**

The scores obtained from the VST and the reading test were used as a proxy to classify subgroups of relatively low- and relatively high-proficiency levels. This procedure involved first ranking all the participants \(N = 256\) in ascending order according to their total VST scores. Once this was established, participants in the top half of this ranked list were flagged as “high,” and those in the bottom half were flagged as “low.” The same procedure was followed again for all the participants according to their reading score. Participants who were in the top 50 percentile for both the vocabulary and reading tests were classified as possessing relatively high overall proficiency, and those who were in the bottom 50 percentile were classified as possessing relatively low proficiency. This resulted in a number of 93 participants being classified as the high-proficiency subgroup and 93 participants as the low-proficiency subgroup.

**Results**

**Learners’ Performance on Vocabulary Levels and Reading Comprehension**

Table 2 presents descriptive statistics of the participants’ scores on the three levels of OVK (OVK1, OVK2, and OVK3) and reading comprehension. As shown in Table 2, the mean score of the performance of the entire cohort in OVK1 was greater than their mean scores on the mid- and low-frequency levels of OVK. The results generally indicated that vocabulary knowledge decreased as word frequency level decreased. The mean score of the reading test, on the other hand, revealed that the group as a whole achieved about 47% of text comprehension.

Descriptive statistics and independent samples \(t\) tests were also conducted to examine the differences between low- and high-proficiency subgroups in terms of their vocabulary knowledge and reading comprehension scores. Results are summarized in Table 3. Independent samples \(t\) tests indicated that there was a statistically significant difference between the mean scores of OVK attained by the relatively low- and high-proficiency participants. This was the case for each of the three frequency-based categories of OVK: OVK1, \(t(180.31) = 36.04, p < .001, Cohen’s d = 3.76\); OVK2, \(t(129.15) = 35.67, p < .001, Cohen’s d = 3.74\); and OVK3, \(t(165) = 13.5, p < .001, Cohen’s d = 1.97\). Independent samples \(t\) tests also indicated a significant difference between the mean reading comprehension test scores of the low- and high-proficiency subgroups, \(t(160.98) = 10.68, p < .001, Cohen’s d = 1.57\).

To compute the contribution of the three levels of OVK (OVK1, OVK2, and OVK3) to reading comprehension, multiple regression analysis was performed. However, before performing multiple regression analysis, correlation coefficients between scores on OVK levels and performance on reading comprehension were examined. All the correlations were strong, positive, and statistically significant (see Table 4).

**Research Question 1:** What is the contribution of OVK of high-, mid-, and low-frequency words in the prediction of L2 reading comprehension for a single cohort of learners?
Prior to running regression analysis, the data were examined for their suitability for such analysis. The standardized predicted values and the residuals indicated that the statistical assumptions for carrying out a regression analysis were met.1

The first phase of the analysis was performed employing a hierarchical multiple regression to determine the degree to which the three levels of OVK (OVK1, OVK2, and OVK3) were capable of explaining the variance in L2 reading comprehension scores for the whole cohort. The first step involved the insertion of OVK1, the high-frequency level, as the independent variable and reading comprehension as the dependent variable. The result indicated a model that could predict about 55% of the variance in L2 reading comprehension. When, in the second step, OVK2 was inputted, a unique explanatory power, accounting for 10.6%, was added to the model, resulting in a model that could explain about 66% of the variance in L2 reading comprehension performance. In the final step, OVK3 was added to the model, but it only provided a marginal non-significant contribution to the overall model (less than 1%). Overall, a model comprising all the three levels of OVK (OVK1, OVK2, and OVK3) could explain about 66% of the variance perceived in L2 reading comprehension, $F(3, 251) = 159.86, p < .001, R^2 = .66$.

The standardized beta weights related to the OVK at steps 2 and 3 of the regression model provide evidence of the relative strength of relationship between OVK1 and OVK2, but not OVK3, and increase within L2 reading comprehension scores. At Step 2 of the analysis, OVK1 ($\beta = 0.32, p < .001$) and OVK2 ($\beta = 0.53, p < .001$) each significantly contributed to increases in L2 reading scores. A comparison of the relative magnitude of the beta weights associated with each level of OVK indicates that mid-frequency words (3,001-9,000), those which are representative of mid-frequency vocabulary, were of a higher relative predictive value to the model (see Table 5).

At Step 3, only OVK1 and OVK2 significantly contributed to the explanatory power of the model. OVK2 remains the strongest predictive variable of the overall predictive capacity of the model ($\beta = 0.56, p < .001$), followed by OVK1 ($\beta = 0.33, p < .001$). OVK3 (9,001-14,000), however, was found to contribute no significant variance in L2 reading comprehension. Thus, OVK at both the

Table 2. Descriptive Statistics of the Learners’ Performance on the Tests (N = 256).

<table>
<thead>
<tr>
<th>Tests</th>
<th>Items</th>
<th>Minimum (%)</th>
<th>Maximum (%)</th>
<th>M (%)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>OVK1</td>
<td>30</td>
<td>50</td>
<td>96.70</td>
<td>69.40</td>
<td>4.03</td>
</tr>
<tr>
<td>OVK2</td>
<td>60</td>
<td>24</td>
<td>58</td>
<td>43.20</td>
<td>6.47</td>
</tr>
<tr>
<td>OVK3</td>
<td>50</td>
<td>2</td>
<td>24</td>
<td>10.07</td>
<td>3.48</td>
</tr>
<tr>
<td>Reading</td>
<td>40</td>
<td>20</td>
<td>72.50</td>
<td>46.85</td>
<td>7.28</td>
</tr>
</tbody>
</table>

Note. OVK = orthographic vocabulary knowledge.

Table 3. Descriptive Statistics of OVK Levels Tests for Lower (n = 93) and Higher Proficiency Subgroups (n = 93).

<table>
<thead>
<tr>
<th>Tests</th>
<th>Lower proficiency</th>
<th>Higher proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (%)</td>
<td>SD</td>
</tr>
<tr>
<td>OVK1</td>
<td>58.53</td>
<td>3.98</td>
</tr>
<tr>
<td>OVK2</td>
<td>32.28</td>
<td>4.38</td>
</tr>
<tr>
<td>OVK3</td>
<td>5.53</td>
<td>2.69</td>
</tr>
<tr>
<td>Reading</td>
<td>31.45</td>
<td>4.32</td>
</tr>
</tbody>
</table>

Note. OVK = orthographic vocabulary knowledge.

Table 4. Correlations Within OVK Levels and Reading Comprehension.

<table>
<thead>
<tr>
<th>Measures</th>
<th>OVK1</th>
<th>OVK2</th>
<th>OVK3</th>
<th>Reading comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>OVK1</td>
<td>—</td>
<td>.77**</td>
<td>.71**</td>
<td>.74**</td>
</tr>
<tr>
<td>OVK2</td>
<td>—</td>
<td>—</td>
<td>.76**</td>
<td>.77**</td>
</tr>
<tr>
<td>OVK3</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.63**</td>
</tr>
</tbody>
</table>

Note. OVK = orthographic vocabulary knowledge.

*Correlation is significant at the level of 0.01 (two-tailed).
high-frequency (0-3,000) and mid-frequency (3,001-9,000) levels contributed significantly to the prediction of L2 reading comprehension performance. This result is reminiscent that for the entire cohort, vocabulary knowledge in the high-frequency range and also knowledge of words in the mid-frequency range were significantly related to L2 reading comprehension results.

The following part of the analysis examines whether the same or different degrees of contribution to L2 reading comprehension by the levels of OVK would emerge when learners’ levels of proficiency were taken into account.

**Research Question 2:** What is the contribution of OVK of high-, mid-, and low-frequency words in the prediction of L2 reading comprehension for learners of relatively low- and relatively high-L2 proficiency levels?

To investigate the relative contribution of OVK of high-, mid-, and low-frequency words in explaining variance within L2 reading comprehension, two additional hierarchical multiple regression models were built. Each of these models was created by including scores from either informants from the relatively low (n = 93) or relatively high (n = 93) proficiency subgroups.

**Results for the low-proficiency subgroup.** From the regression model summarized in Table 6, it can be observed that only OVK1 (0-3,000-word level) provided a unique contribution to L2 reading comprehension performance for the low-proficiency subgroup. Step 1 of the hierarchical multiple regression analysis, which only included OVK1, was found to explain about 18% of the variance in L2 reading comprehension scores among the learners with relatively low-proficiency level, $F(1, 91) = 19.83, p < .001, R^2 = .18$. When OVK2 was added in Step 2, an additional 4% contributed to the overall explanatory power of L2 reading performance, resulting in a model that could explain about 22% of the variance, $F(2, 90) = 12.76, p < .001, R^2 = .22$. However, although significant, the contribution of OVK2 (mid-frequency vocabulary) to reading comprehension of low-proficiency learners is small when compared with that of OVK1. Adding OVK3 to the regression model did not appear to contribute any additional predictive value of L2 reading comprehension at this proficiency level. Therefore, it is suggested that OVK1 has the greatest contribution to explaining reading comprehension in L2 learners of a relatively low-proficiency level, whereas Level 2 has only a small effect.

**Results for the high-proficiency subgroup.** The regression model, summarized in Table 7, indicates that when only OVK1 (Step 1) was entered, it explained about 18% of the variance in L2 reading comprehension scores, $F(1, 91) = 19.29, p < .001, R^2 = .18$. At Step 2, both OVK1 and OVK2 were included. OVK2 was found to contribute an additional 26% of the variance in reading comprehension. When combined, these two levels explained about 43% of the variance, $F(2, 90) = 33.96, p < .001, R^2 = .43$, lending support to a greater explanatory power of L2 reading comprehension model in high-proficiency learners when the mid-frequency variable was factored in. At Step 3, which comprised all the three levels, the regression model strength has slightly increased, $F(3, 89) = 27.81, p < .001, R^2 = .48$, adding about 5% that was explained by OVK3 (low-frequency words).

Appraisal of the beta weights associated with each of the OVK levels in the last step of the model (see Table 7) indicates that all the three levels contributed significantly to the predictive power of L2 reading comprehension. However, a comparison of the relative magnitude of the beta weights

---

**Table 5. Regression Model Including OVK Levels for the Total Group (N = 256).**

<table>
<thead>
<tr>
<th>Step</th>
<th>$R$</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
<th>Unstandardized</th>
<th>Standardized</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$B$</td>
<td>$\beta$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$SEB$</td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>.74</td>
<td>.55***</td>
<td></td>
<td>-9.24</td>
<td>1.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.35</td>
<td>.74***</td>
</tr>
<tr>
<td>Step 2</td>
<td>.81</td>
<td>.66***</td>
<td>.11***</td>
<td>-5.70</td>
<td>1.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.59</td>
<td>.32***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.57</td>
<td>.53***</td>
</tr>
<tr>
<td>Step 3</td>
<td>.81</td>
<td>.66***</td>
<td>0.00</td>
<td>-6.14</td>
<td>1.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.61</td>
<td>.33***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.61</td>
<td>.56***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.11</td>
<td>-.05</td>
</tr>
</tbody>
</table>

Note. Numbers are rounded to two decimal points. OVK = orthographic vocabulary knowledge; SEB = standard error beta. $^{***}p < .001$. 

---
attributed to Step 3 of the model suggests that levels 1 and 2, which are representative of words of high- and mid-frequency levels, are of a larger relative predictive value to the model of L2 reading comprehension than those of OVK3, the low-frequency words. A synthesis of the key findings from the analyses performed in the current study is presented in Table 8.

### Discussion

The present study examined the relative contribution of the high-, mid-, and low-frequency vocabulary to reading comprehension among L2 learners. Two research questions were addressed in the study in relation to this objective. Research Question 1 addressed the contribution of the three levels of vocabulary knowledge to reading comprehension in relation to a single cohort of learners. Results showed that for the entire cohort involved in the study, OVK of words across high- and mid-frequency ranges was collectively capable to predict about 66% of the variance observed in L2 reading comprehension scores. This finding, coupled with the strong correlations observed between OVK and L2 reading comprehension, suggests that OVK of L2 words from the high- and mid-frequency ranges was strongly related to performance on L2 reading comprehension.

The strength of connection observed between the measures of written receptive vocabulary knowledge and L2 reading comprehension scores \((r = .63-.77, p < .001)\) was similar to

<table>
<thead>
<tr>
<th>Table 6. Regression Model Including OVK Levels for the Low-Proficiency Group ((n = 93)).</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Step 1</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>OVK1</td>
</tr>
<tr>
<td>Step 2</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>OVK1</td>
</tr>
<tr>
<td>OVK2</td>
</tr>
<tr>
<td>Step 3</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>OVK1</td>
</tr>
<tr>
<td>OVK2</td>
</tr>
<tr>
<td>OVK3</td>
</tr>
</tbody>
</table>

**Note.** OVK = orthographic vocabulary knowledge; SEB = standard error beta. 
* ***p < .05. ** ***p < .001.

<table>
<thead>
<tr>
<th>Table 7. Regression Model Including OVK Levels for the High-Proficiency Group ((n = 93)).</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Step 1</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>OVK1</td>
</tr>
<tr>
<td>Step 2</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>OVK1</td>
</tr>
<tr>
<td>OVK2</td>
</tr>
<tr>
<td>OVK3</td>
</tr>
</tbody>
</table>

**Note.** OVK = orthographic vocabulary knowledge; SEB = standard error beta. 
* ***p < .01. ** ***p < .001.
those reported in previous related studies. For example, Stehr (2008) reported a correlation of .83 between reading comprehension and measures of written receptive vocabulary knowledge. Similarly, Milton et al. (2010) reported that OVK positively correlated with reading comprehension, as measured by the reading subsection of IELTS ($r = .70, p < .01$), and was found to explain about 50% of the variance in L2 reading comprehension. The capacity of the scores on OVK measures used in the present study to predict variance in L2 reading comprehension ($R^2 = .66$) was slightly larger than that observed in Milton et al.’s (2010) study. The best interpretation of this is that the vocabulary test used in the current study was capable of measuring word knowledge beyond the 5,000-word limit test used in Milton et al.’s (2010) study. However, the main focus of the current study was to disentangle the predictive power of each frequency level, high, mid, and low, in L2 reading comprehension. For the entire single cohort of learners, both high- and mid-frequency vocabulary were observed to contribute uniquely to L2 reading comprehension. Low-frequency vocabulary, however, did not appear to explain any additional value in the predictive model of L2 reading performance.

Research Question 2 sought to determine the degree to which high-, mid-, and low-frequency vocabulary levels were proportionally predictive of L2 reading comprehension scores for learners of relatively low and relatively high overall proficiency. In the learners of relatively low-proficiency subgroup, only OVK1 and OVK2 were able to predict variance in L2 reading comprehension. The greatest predictive value was explained by the knowledge of high-frequency vocabulary ($R^2 = .18$), followed by the mid-frequency vocabulary ($R^2 = .04$). Similar to the reading model established for the entire cohort of learners, the low-frequency words contributed no additional value in predicting L2 reading comprehension. This finding suggests that the low level of knowledge of vocabulary beyond the 3,000 most frequent words in the relatively low-proficiency subgroup was not sufficient to impart a measurable effect on their L2 reading comprehension achievement.

For the relatively high-proficiency subgroup, the first step of the regression model that included the entry of only OVK1, high-frequency vocabulary, showed a similar predictive value of high-frequency words in reading (18%) as that found in the low-proficiency subgroup. However, when OVK2 (mid-frequency vocabulary) was inputted in the second step of the regression with OVK1, an additional 26% of the variance in reading was explained by mid-frequency vocabulary. This finding lends support to the efficacy and usefulness of the mid-frequency words, which are representative of the word frequency range of 3,001 to 9,000, in explaining a unique variance in L2 reading comprehension, over and above high-frequency words range, within the relatively high-proficiency subgroup. The contribution of mid-frequency words to L2 reading comprehension found in this study supports the findings from earlier research. For example, Laufer and Ravenhorst-Kalovski (2010) found that knowledge of the 6,000 to 8,000 word families was required for their students to achieve 98% of text comprehension. Also, Schmitt and Schmitt (2014) argued that even the ability to read with some guidance and support requires 95% of text coverage, which entails knowledge of 4,000 to 5,000 word families. Therefore, even assisted reading in an educational setting requires a substantial progression into mid-frequency vocabulary. The final step of the regression analysis in this subgroup included all the three levels of OVK (OVK1, OVK2, and OVK3). The result suggests that OVK3, low-frequency range (9,001-14,000), only added a marginal effect to the performance in reading comprehension scores. This finding is in line with that from previous studies (e.g., Nation, 2006; Schmitt et al., 2011), where vocabulary knowledge of 8,000 to 9,000 word families would provide learners with adequate comprehension. Vocabulary situated in frequency range beyond the 9,000-word level appeared to contribute only a very small value to the level of reading comprehension that does not merit the burden taken to learning those words.

Overall, the findings of the present study suggest a strong association between OVK of words beyond the most frequent 3,000 word families, and the L2 reading performance in learners with a relatively high overall L2 proficiency. This finding is, thus, suggestive of the benefit of L2 learners possessing robust OVK of words in the mid-frequency range, over and above knowledge of only high-frequency words.

**Pedagogical Implications**

This study provides evidence that can inform pedagogical approaches to the development of L2 vocabulary knowledge.
for the purpose of supporting L2 reading comprehension. In the literature devoted to the relationship between vocabulary knowledge and reading comprehension, we have seen the benefits of developing a relatively large total vocabulary size, but in fact the three different frequency levels (i.e., high, mid, and low) have been treated quite differently in language teaching (Schmitt & Schmitt, 2014). This reason might be enough to study the effect of each frequency level on reading comprehension to provide language teachers, learners, and practitioners with useful pedagogical implications.

First, although knowledge of the most frequent 3,000 word families per se cannot be argued to provide a sufficient level of reading comprehension, it appears to correlate significantly with reading ability and explain unique variance in this construct. This finding along with the text coverage (about 82%) these frequent words provides supports a recommendation for L2 learners to have a good level of mastery of these words (Nation, 2001; Schmitt & Schmitt, 2014). The suggestion here is that instead of being considered a volume of vocabulary knowledge that facilitates adequate L2 reading achievement, knowledge of the high-frequency vocabulary is considered in this study as a step toward developing a larger vocabulary size based on which, through learning process, remarkable levels of L2 reading comprehension may be established. As Nation (2001) has pointed out, any effort to have these words learned should be made because they provide learners with the knowledge to become independent, at least to a limited extent, and direct their own vocabulary learning. One way to achieve this is via extensive reading and the reading of easy and enjoyable materials in large quantities (Day & Bamford, 2002; Masrai & Milton, 2018a). Extensive reading allows learners to engage with words in a contextualized and authentic environment, giving a rich of a meaningful input. Furthermore, in addition to the ongoing development of vocabulary, pedagogical emphasis on the development of reading strategies is also warranted. Since learners at a low-proficiency level are unlikely to have sufficient vocabulary knowledge to handle the processing demands of L2 reading, deliberate instruction on the effective operationalization of reading strategies also needs to be delivered while the OVK of words from the high-frequency range (0-3,000) is being developed.

Second, concerning the value of developing OVK beyond the high-frequency range, the significance of mid-frequency words to predict reading comprehension for the participants of the relatively high-proficiency subgroup cannot be overstated. In this group of learners, mid-frequency vocabulary contributed greatly to the reading comprehension model. Thus, it could be emphasized here that although knowledge of high-frequency vocabulary may provide learners with a large coverage of running words they are likely to encounter while reading authentic materials, the expansion of a learner’s OVK to include the mid-frequency words range is most likely to have a notable impact on the speed of lexical access when reading (Laufner & Nation, 2001; Schmitt & Schmitt, 2014). Despite the considerable importance and benefits of mid-frequency vocabulary, at least among learners at or near C1 level, which the present study confirms, it is not often addressed pedagogically (Schmitt & Schmitt, 2014). It can be argued that some language teachers might assume that vocabulary will somehow be learned from exposure to different language activities delivered within the classroom time and from informal input outside the classroom. There is evidence, however, that mid-frequency words are not typically used or taught by teachers in classrooms to any large extent. Horst, Collins, and Cardoso (2009, cited in Schmitt & Schmitt, 2014), for example, found that the overwhelming majority of cases of direct vocabulary teaching in language classrooms focused on high-frequency words, leaving very little focus on mid-frequency words. A number of studies (e.g., Horst, 2010; Tang & Nesi, 2003) also analyzed data from teachers’ talk in language classrooms and found that only a small proportion of vocabulary in the mid-frequency range were used by teachers. Furthermore, these studies showed that even the very little quantity of mid-frequency vocabulary used by teachers did not receive the required number of repetitions to facilitate its acquisition.

Evidence of less attention to mid-frequency vocabulary also came from the studies which have examined EFL textbooks. For example, Matsuoka and Hirsh (2010) analyzed vocabulary presented in New Headway: Upper-Intermediate English textbooks and found that most of the words appeared in these textbooks are from the high-frequency range, and those words beyond this level were very minimally recycled, only one to five times. Having addressed the importance of mid-frequency vocabulary and the issues surrounding its teaching, it is highly recommended that words from this frequency range should receive considerable attention by language teachers, textbook writers, and policy-makers in language learning. Programs which are aimed at supporting L2 reading comprehension through vocabulary development should place a pedagogical focus on increasing learners’ engagement with mid-frequency vocabulary in the orthographic modality. Furthermore, diagnostic and formative assessment of OVK of words in this frequency range should be performed regularly to ensure their learning.

Finally, developing a vocabulary knowledge of 8,000-word families is the important criterion for attaining the kind of fluency associated with higher-level study through the medium of English. Studies such as Masrai and Milton (2018b) and Townsend, Filippini, Collins, and Biancarosa (2012) showed that educational performance is greatly predicted by a large general vocabulary size than by a narrower vocabulary even when highly specialized.

Limitations

There are some limitations to the present study which will be addressed to help future research in this area. One limitation is that although the VST used in the current study was useful,
to some extent, in providing scores related to each frequency level, the low sampling rate of words in each frequency level might not be representative of words knowledge in each frequency level. As there is no existing VLT, to the researcher’s knowledge, that covers high-, mid-, and low-frequency vocabulary with a high sampling rate, devising a test that has this feature is rather needed to clearly pinpoint the contribution of these different frequency levels to L2 learners’ reading comprehension. While this is a standardized test intended to predict ability to read when studying academic content through the medium of English, the task-type and topic familiarity are expected to affect reading comprehension of the participants in the present study. Finally, the study was conducted within a relatively homogeneous informant group in accordance with background demographics. Thus, conducting such a study among language learners of different L1 backgrounds, a wide range of proficiency levels, ages, and educational levels would provide more generalizable findings.

Conclusion

This study has addressed the importance of vocabulary knowledge in L2 reading comprehension. Both high- and mid-frequency vocabulary were found to contribute unique variance in explaining reading comprehension among L2 learners. The contribution of high- and mid-frequency vocabulary to L2 reading comprehension differs depending on the overall proficiency level of learners. In the relatively low-proficiency subgroup, high-frequency vocabulary was the only predictor of reading comprehension. In the relatively high-proficiency subgroup, on the other hand, both high- and mid-frequency vocabulary significantly contributed to reading comprehension, where mid-frequency vocabulary explained the largest variance. Based on this evidence, it can be suggested that the development of L2 OVK beyond the high-frequency range would greatly enhance L2 learners’ reading performance.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. An analysis of the multicollinearity statistics showed that the tolerance values (0.36, 0.31, and 0.33) of three independent variables (OVK1, OVK2, and OVK3) were all higher than 0.1 and variance inflation factor (VIF) values (2.77, 3.41, and 3.07) were lower than 10; therefore, there was no problem of multicollinearity among independent variables (tolerance should be higher than 0.1 and VIF should be lower than 10 according to Field, 2013, p. 325).

Supplemental Material

Supplemental material for this article is available online.

ORCID iD

Ahmed Masrai https://orcid.org/0000-0002-6778-5952

References


Staehr, L. (2008). Vocabulary size and the skills of listening, reading and writing. Language Learning, 36, 139-152.


Author Biography

Ahmed Masrai is an assistant professor of Applied Linguistics in the Department of Languages at King Abdulaziz Academy, Saudi Arabia. His research and teaching interests include second language acquisition, lexical studies, and vocabulary testing. He has published in a number of journals, including Journal of English for Academic Purposes, Applied Linguistics Review, Language Assessment Quarterly and Language Learning Journal.