The Contribution of Vocabulary Knowledge to Reading Proficiency: The Case of College Russian

Jane F. Hacking
University of Utah

Erwin Tschirner
Universität Leipzig

Abstract: Literacy development in a second language (L2) is a key goal of college foreign language study; language programs aspire to graduate students with strong L2 reading ability. Research has shown a strong correlation between L2 vocabulary knowledge and L2 reading proficiency; however, much of this research has focused on English as a second language. Research among other language learning populations is needed to offer empirically grounded suggestions on how best to achieve high levels of L2 literacy. This article presents data on the reading proficiency and vocabulary knowledge of 48 native American-English-speaking college-level learners of Russian in order to determine whether there are identifiable lexical thresholds associated with moving from one level of reading proficiency to the next. Participants completed the ACTFL Reading Proficiency Test rated according to the ACTFL Proficiency Guidelines 2012—Reading and the Russian Vocabulary Levels Test, which measured knowledge of the 5,000 most frequent words in Russian. Results showed that there are statistically significant lexical minimums associated with different levels of reading proficiency. These findings suggest the utility of similarly designed studies for other languages and are discussed in terms of implications for the role of developing vocabulary knowledge in an undergraduate curriculum.

Key words: Russian, postsecondary/higher education, proficiency, program goals and development, vocabulary
Literacy development in a second language \( (L2) \) is a key goal of college foreign language study. This objective is succinctly captured in the 2009 Report to the Teagle Foundation on the Undergraduate Major in Language and Literature, prepared by the Modern Language Association, which states that “\( R \)eaching advanced literacy and linguistic levels should be the expected outcome for all language majors” (Modern Language Association, 2009, n.p.). The report affirms the “centrality of literature and reading to undergraduate education.” There has been substantial research investigating the role that vocabulary knowledge plays in the development of overall L2 proficiency (see Milton, 2009, for an overview), and a subset of this research on vocabulary and L2 proficiency has focused specifically on the question of reading ability (e.g., Nation, 2006; Tschirner, 2004). Intuitively, one expects that as L2 proficiency increases so too does vocabulary size, and indeed this has been shown to be the case (e.g., Alderson, 2005; Staehr, 2008). Much of the empirical research to date has focused on learners of English (Nation, 2013; Schmitt & Schmitt, 2014; but see David, 2008, for a study of French learners) and in particular on considering what level of English vocabulary knowledge learners require to be successful in different contexts, e.g., how much vocabulary is needed to be conversant in informal settings. What vocabulary level is required to matriculate into a degree program instructed in an L2? This article presents data on the reading proficiency and vocabulary knowledge of 48 native American-English-speaking college-level learners of Russian and addresses three questions. First, do the data for Russian learners also evidence the strong correlation between vocabulary knowledge and reading proficiency that has been documented for learners of English? Second, is it possible to identify lexical thresholds associated with moving from one level of proficiency to the next? And third, how might these data help language educators think about undergraduate curricular goals? Specifically, what do these findings suggest about the role of developing vocabulary knowledge in the undergraduate Russian curriculum? The immediate focus of this research is the relationship between vocabulary knowledge and reading proficiency for a group of L2 Russian learners. The discussion of the curricular and pedagogical implications of this relationship will, however, resonate with instructors of other languages who share the goal of developing L2 literacy in undergraduate language programs.

**Background**

**Vocabulary Size and L2 Development**

Existing L2 learner vocabulary studies offer robust findings documenting the critical role that vocabulary plays in L2 acquisition. For example, Alderson (2005, n.p.) reported data showing that vocabulary knowledge accounts for 37–62% of the variance in language proficiency scores, and Milton (2013, p. 67) concluded that correlations between receptive vocabulary tests and reading proficiency ranged from 0.50 to 0.85, explaining between 25 and 72% of the variance. Proceeding from this general premise of the centrality of vocabulary knowledge to L2 development, L2 vocabulary research has focused on a variety of specific questions, such as the following: How do learners acquire L2 vocabulary? How can language educators effectively measure vocabulary knowledge? Is it possible to establish reliable lexical thresholds—that is, estimates of the minimum number of lexical items that are required to achieve a certain result (e.g., a particular proficiency rating, exam score, or reading comprehension score)? The concern here is primarily with this final question as we consider the relationship between students’ levels of Russian reading proficiency and their Russian vocabulary knowledge.

**Measuring Vocabulary Size**

There are two dimensions of vocabulary knowledge that researchers try to measure:
size and depth. Measurements of vocabulary size typically seek to count the number of words a learner knows in a basic form-meaning way, whereas depth measures attempt to capture a learner’s grasp of such things as a word’s possible collocations, associations, and constraints on use. A second important parameter in vocabulary studies is whether it is receptive or productive knowledge that is being measured. Receptive vocabulary knowledge is typically associated with reading and listening, while productive knowledge is demonstrated through speaking and writing. Despite suggestions that the two types of knowledge may necessarily interact—for example, (1) being a good listener (receptive) may involve being able to anticipate what is coming (productive) (Milton, 2009, p. 13), or (2) models of learning that view L2 vocabulary development as moving from receptive to productive capacity—the field tends to view receptive and productive vocabulary knowledge separately, especially for the purposes of testing. Since the emphasis here is on reading, the study focused on receptive vocabulary knowledge.

Tests of receptive vocabulary knowledge are based on the idea of word frequency. It is generally posited that more frequent words are learned (or should be learned) earlier (Milton, 2007). Tests, including the one used in this study, assess frequency bands in 1,000-word increments. Test takers are exposed first to a sample of the 1,000 most frequent words and receive a score for the number correct. They then move on to the next band containing words sampled from the 1,001–2,000 most frequent words, and so on. The expectation is that a learner will know fewer and fewer words as she or he moves from the most frequent to increasingly less frequent words.

**Lexical Thresholds**
Much vocabulary research is concerned with the question of lexical thresholds, or how much vocabulary knowledge is required to achieve certain goals. For example, Schmitt and Schmitt (2014) asserted that for English learners, an optimal threshold of 9,000 words allows L2 readers to read a wide range of authentic texts in English without being slowed down substantially by unknown vocabulary. In contrast, knowledge of only the most frequent 3,000 words allows learners to derive some meaning from authentic texts, read graded readers with full understanding, and understand dictionary definitions (Cobb, 2007; Schmitt & Schmitt, 2014).

Another approach is to peg vocabulary size to particular levels on standardized tests of L2 proficiency. Milton (2010) correlated the results of the XLex vocabulary test (Meara & Milton, 2003) with levels of the Common European Framework of Reference (CEFR) for learners studying English, French, and Greek as a foreign language and suggested that vocabulary knowledge of around 3,000 words is required for learners to reach Level B1 and 5,000 for Level C1. He also noted that there seemed to be differences among languages. In French, for example, it seemed that fewer words were required at a particular CEFR level than in English, while in Greek, more words seemed to be required than in English. Whereas an average of 2,500 words appeared to be sufficient to be rated at Level B1 in French, an average of 3,500 words appeared to be required to be at the same overall level in Greek. Meara and Milton (2003, cited in Milton, 2010, p. 224) argued that 1,500–2,500 words are required for Level A2, 2,750–3,250 for Level B1, 3,250–3,750 for Level B2, 3,750–4,500 for Level C1, and 4,500–5,000 for Level C2 in English as a second language (ESL). Huhta, Alderson, Nieminen, and Ullakonoja (2011, n.p.) compared the results of Nation’s Vocabulary Levels Test (Nation, 1990) and the DIALANG English reading tests (Alderson & Huhta, 2005) and found that the 2,000-word level test distinguished between CEFR Levels A2 and B1, the 3,000-word test between B1 and B2, and the 5,000-word test between B2 and C1.
Vocabulary Knowledge and L2 Reading Comprehension

The notion of lexical thresholds is closely associated with text coverage—that is, the percentage of a text that a reader can be expected to understand. The consensus among researchers has been that a learner needs to know between 95 and 98% of a text’s running words in order to comprehend the text (Carver, 1994; Hirsh & Nation, 1992; Hu & Nation, 2000; Nation, 2006; Schmitt, 2008; Schmitt, Jiang, & Grabe, 2011). For English learners, research has shown that the first 9,000 word families provide coverage of 98% of words in a wide range of texts (Nation, 2006). Clearly, unsimplified text is not readily accessible for L2 learners until they have acquired substantial vocabulary knowledge. As Nation (2007) noted, “most text beyond the 3,000-word level of graded readers series is very difficult for foreign language learners. This is because in most novels a very large number of different words occur beyond the learners’ current vocabulary knowledge” (p. 9). Of prime concern for both researchers and instructors is how to negotiate the challenge that limited vocabulary knowledge presents, particularly because current language pedagogy is committed to the use of authentic, i.e., unsimplified, texts, a position that is consistent with the goal outlined in the Teagle Report of graduating majors who have reached “advanced literacy and linguistic levels” (Modern Language Association, 2009, n.p.).

Vocabulary Knowledge and L2 Russian

The field of L2 Russian does not have the extensive body of research on the contribution of vocabulary knowledge to proficiency development that exists for English learners. However, lexical minimums are laid out for each level of the Test of Russian as a Foreign Language (TORFL)/Тест по русскому языку как иностранный (ТРКИ) developed by the Russian Ministry of Education and Science (http://russian-test.com/tests/torfl). Some official test specifications can be found on the Web site of Moscow State University’s Training and Testing Language Center for Foreigners (http://russian-test.com/tests/torfl/). These standardized tests of Russian language proficiency were designed as part of the CEFR for languages; learners may achieve one of six levels: Elementary, Basic, Level 1, Level 2, Level 3, and Level 4. Each level outlines a set of competencies and is accompanied by a description of what language at that level enables the learner to do. For example, a learner who scores at the Basic Level, the level required to become a naturalized Russian citizen, is described as being able to “satisfy the most basic communicative needs . . . in a limited number of predictable situations” (http://russian-test.com/assets/docs/Trebovaniya_-_basic.pdf). Official documentation for the test specifies a minimum number of vocabulary words required at a given level as shown in Table 1. Beginning with Level 2, vocabulary knowledge is explicitly split into receptive and productive vocabulary with the expectation that learners’ productive vocabulary knowledge will constitute a subset of their larger receptive lexicon. Test documentation states that a learner must have achieved Level 1 to begin a course of study at a Russian institution of higher education and that Level 2 proficiency is necessary to receive a degree taught in Russian (with the exception of degrees in philology, for example, for which Level 3 proficiency is the stated requirement).

The TORFL outlines the following correspondences between levels on the TORFL and levels on other tests, as shown in our compilation of standard comparisons of these tests in Table 2. It is unclear if the suggested correspondences were based on empirical studies. While the correspondences between the CEFR, the ACTFL, and the International Language Roundtable (ILR) seem to be fairly consistent with established crosswalks (ACTFL, 2016), the correspondence between the TORFL Level
4 and the ILR 4/4+ does not appear to be realistic. The ACTFL Superior level corresponds to the ILR 3/3+ level and not the ILR 4/4+ level. In addition, the suggested correspondences between TORFL levels and vocabulary sizes (compare with Table 1) are very different from the ones established by previous research. In this study, we examine which vocabulary sizes correspond with which ACTFL reading proficiency levels to provide initial evidence for two hypotheses: (1) There is a positive correlation between reading proficiency and vocabulary size for Russian; and (2) The correspondences between reading proficiency and vocabulary size are consistent with what is known about other Indo-European languages.

Because there are two established mastery criteria in the literature investigating the relationship between reading proficiency and vocabulary size (80% correct, e.g., Xing & Fulcher, 2007, or 85% correct, e.g., Schmitt, Schmitt, & Clapham, 2001), the study addressed the following research question first:

1. Which of the two mastery criteria provides a better description of the relationship between reading scores and vocabulary size measures?

The study then separately examined reading proficiency and vocabulary sizes to answer the following questions:

2. What are the average reading proficiency levels of students at various semester levels studying Russian at a large Western state university in the United States?

### Table 1

<table>
<thead>
<tr>
<th>Proficiency Level</th>
<th>Minimum Vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary</td>
<td>780</td>
</tr>
<tr>
<td>Basic</td>
<td>1,300</td>
</tr>
<tr>
<td>Level 1</td>
<td>2,300</td>
</tr>
<tr>
<td>Level 2</td>
<td>10,000 (6,000 in active vocabulary)</td>
</tr>
<tr>
<td>Level 3</td>
<td>12,000 (7,000 in active vocabulary)</td>
</tr>
<tr>
<td>Level 4</td>
<td>20,000 (8,000 in active vocabulary)</td>
</tr>
</tbody>
</table>

*Note: Based on information provided at http://russian-test.com/tests/torfl/*

### Table 2

<table>
<thead>
<tr>
<th>Instrument/Rating Scale</th>
<th>TORFL</th>
<th>CEFR</th>
<th>ACTFL</th>
<th>ILR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary</td>
<td></td>
<td>A1</td>
<td>Novice</td>
<td>0/0+</td>
</tr>
<tr>
<td>Basic</td>
<td></td>
<td>A2</td>
<td>Intermediate Mid</td>
<td>1</td>
</tr>
<tr>
<td>Level 1</td>
<td></td>
<td>B1</td>
<td>Intermediate High</td>
<td>1+</td>
</tr>
<tr>
<td>Level 2</td>
<td></td>
<td>B2</td>
<td>Advanced Mid</td>
<td>2/2+</td>
</tr>
<tr>
<td>Level 3</td>
<td></td>
<td>C1</td>
<td>Advanced High</td>
<td>3</td>
</tr>
<tr>
<td>Level 4</td>
<td></td>
<td>C2</td>
<td>Superior</td>
<td>4/4+</td>
</tr>
</tbody>
</table>
3. What are the average vocabulary sizes of students at various semester levels studying at a large Western state university in the United States?

To inform an understanding of the relationship between reading proficiency and vocabulary knowledge, the study posed two final questions:

4. What is the correlation between the reading proficiencies and the vocabulary sizes of U.S. college students studying Russian?
5. Which vocabulary sizes correspond to which reading proficiency levels of U.S. college students studying Russian?

Methods

Participants
The 48 participants in this study were college students of Russian at a large U.S. state university. The distribution of students across years of study was 31.3% first-year students, 14.6% second-year students, and 54.2% upper-division students. Of all participating students, 50% had previously spent 18–24 months in a Russian-speaking country, while 39.6% of the participants were female and 60.4% were male.

Instruments and Procedures
Data were gathered using two instruments, described below, at two points during the academic year. First- and second-year students were tested at the end of their first or second year of coursework. In the upper divisions, testing was conducted at different points depending on the student profile. The majority of the participants who were tested enrolled in upper-division courses following a period of 18–24 months of service in a Russian-speaking country. Prior to departure, they would have received several weeks of intensive language instruction focused primarily on achieving the goals of their service mission. These students were tested upon returning to the United States at the beginning of their third-year course, i.e., in fall 2015. Two additional upper-division students who did not have experience abroad (one third-year student and one graduating senior) were tested at the end of the academic year, i.e., in spring 2016. Data for the upper-division students were combined across testing dates and are reported together. As a whole, 47.9% of the students took the ACTFL Reading Proficiency Test (RPT; ACTFL, 2013) and the Russian Vocabulary Levels Test (RVLT; http://www.itt-leipzig.de/static/startseiteeng.html) at the beginning of the fall 2015 semester and 52.1% took the tests at the end of the spring 2016 semester.

The RPT
The RPT is a standardized global assessment of reading ability in a language (ACTFL, 2013). The test measures how well a person spontaneously reads texts and successfully carries out the text-related comprehension tasks that are described at each proficiency level in the ACTFL Proficiency Guidelines from 2012 (ACTFL, 2012). For example, an Intermediate task requires understanding information that is contained in one sentence, whereas Advanced tasks require the ability to understand information that is spread out over several sentences or paragraphs. Each text is accompanied by three multiple-choice items with four possible responses; only one is correct. Both the tasks and multiple-choice responses are presented in the target language (Tschirner, 2016).

The RPT is a timed test with a total test time of 25 minutes allocated to items at each sublevel. Test takers usually complete the readings for two to three sublevels at the same testing session. If a test taker completes the readings for two sublevels, then those responses are considered together to determine the final rating. If a participant completes reading passages for more than two sublevels, the two highest sublevels that can be rated are considered. For the present study, participants were presented with 10–15 texts and their accompanying
multiple-choice items at the Intermediate Low (IL), Intermediate Mid (IM), Advanced Low (AL), Advanced Mid (AM), or Superior sublevels. Because there are no Novice texts and tasks, the three Novice sublevels are determined according to how close the test taker is to the Intermediate level. Because the test is semi-adaptive, it first presents passages at the Advanced level. Depending on how the student copes with the initial items, the test then moves down to the Intermediate or up to the Superior level. The algorithm and cut points that are used to determine a rating at each level were determined empirically (Institute for Test Research and Test Development, 2013). For 54.2% of the students, the final rating was established after 15 texts and 45 items; for the remaining 45.8% of the students, it was established after 10 texts and 30 items. The test is Internet-administered and computer-scored (ACTFL, 2013).

The RVLT
The RVLT consists of a receptive and a productive test (Institute for Test Research and Test Development, n.d.). It is modeled after the English Vocabulary Levels Test pioneered by Nation (1990). The RVLT measures how many of the most frequent 5,000 words of Russian (Sharoff, Umanskaya, & Wilson, 2013) are known. It consists of five bands: the most frequent 1,000; 1,001–2,000; 2,001–3,000; 3,001–4,000; and 4,001–5,000 words. The receptive test, which was used in the present study, consists of 10 clusters of six words each for each of these five bands. Each band is thus represented by 60 words. These words consist of 30 nouns, 18 verbs, and 12 adjectives and are chosen at random from the 1,000 words that constitute each band. Each cluster focuses on one part of speech. Three words of a cluster are targets, which need to be defined by choosing from a list of synonyms and paraphrases. The other three words are distractors. The definition of receptive mastery of a particular band varies slightly in the literature. The two most common percentages used are 80% (e.g., Xing & Fulcher, 2007) and 85% (e.g., Schmitt et al., 2001).

Data Coding
Following Rifkin (2005) and others, the ACTFL RPT results were coded numerically as follows: Novice Low (NL) = 1, Novice Mid (NM) = 2, Novice High (NH) = 3, IL = 4, and so on, up to Superior = 10. The RVLT results were analyzed to determine if the words of a particular band, e.g., the most frequent 1,000 words, were known. The two mastery criteria that have been most commonly used in other studies—80% correct and 85% correct, mentioned above—were used. The highest band at which students got 80 or 85% correct was considered their vocabulary level.

Results

Reading Proficiency
Participants took between 21 and 63 minutes to complete the RPT. The mean time for the two-sublevel test (N = 22; maximum time allowed = 50 minutes) was 33.5 minutes (SD = 6.08), the minimum was 21 minutes, and the maximum was 45 minutes. The mean time for the three-sublevel test (N = 26; maximum time allowed = 75 minutes) was 50.6 minutes (SD = 7.73), the minimum was 38 minutes, and the maximum was 63 minutes.

The results of the RPT ranged from NL to Superior, as shown in Table 3. Nineteen students were rated Novice, fifteen were Intermediate, seven were Advanced, and seven were Superior. Thus, close to 40% were rated Novice, approximately 30% were rated Intermediate, and another 30% were rated Advanced or Superior.

To facilitate a discussion of the development of reading proficiency and its relationship to vocabulary knowledge as students move through an undergraduate program, the minimum, the maximum, the mean, the 95% confidence level of the mean, and the standard deviation of RPT scores by class level are provided in Table 4.
The average proficiency of first-year students was NM (1.73), the average proficiency of second-year students was NH (3.29), and the average proficiency of upper-division students was AL (6.65).

**Vocabulary Scores**

The maximum time allowed for the RVLT was 25 minutes, approximately 5 minutes per band. Within the allocated time block, however, test takers could spend as much time on any single band as they wanted. To determine the mastery criterion for each band, a composite score consisting of the summed individual band scores was calculated and correlated with both mastery criteria: 80% and 85%. The maximum possible score for the five bands was 150. The actual maximum score achieved by a student was 146 and the minimum was 8. The mean was 70.25 and the standard deviation was 45.86. Pearson’s $r$ correlation between the composite vocabulary score and the vocabulary levels as calculated on the basis of the 80% criterion was 0.96 ($p < 0.01$), while for the 85% criterion it was 0.92 ($p < 0.01$). The 80% criterion therefore was chosen as the mastery criterion because it yielded a higher correlation with the total number of items correct than did the 85% criterion. Table 5 shows the distribution of the

---

**TABLE 3**

<table>
<thead>
<tr>
<th></th>
<th>ACTFL</th>
<th>NL</th>
<th>NM</th>
<th>NH</th>
<th>IL</th>
<th>IM</th>
<th>IH</th>
<th>AL</th>
<th>AM</th>
<th>AH</th>
<th>S</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>13</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>7</td>
<td>48</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 4**

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>15</td>
<td>1</td>
<td>5</td>
<td>1.73</td>
<td>0.99</td>
<td>2.47</td>
<td>1.34</td>
</tr>
<tr>
<td>Second</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>3.29</td>
<td>1.80</td>
<td>4.77</td>
<td>1.60</td>
</tr>
<tr>
<td>Upper-Division</td>
<td>26</td>
<td>1</td>
<td>10</td>
<td>6.65</td>
<td>5.59</td>
<td>7.71</td>
<td>2.66</td>
</tr>
<tr>
<td>Total</td>
<td>48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 5**

<table>
<thead>
<tr>
<th>Level</th>
<th>0</th>
<th>1,000</th>
<th>2,000</th>
<th>3,000</th>
<th>4,000</th>
<th>5,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>25</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Percent</td>
<td>52.1</td>
<td>10.4</td>
<td>6.3</td>
<td>6.3</td>
<td>10.4</td>
<td>14.6</td>
</tr>
</tbody>
</table>
vocabulary results. Note that the results reflect the highest band on which a test taker scored at least 80% correct.

Table 5 shows that approximately half of the students did not yet have receptive mastery of the first 1,000 words of Russian while approximately 23% had mastered between 1,000 and 3,000 words and 25% had mastered between 4,000 and 5,000 words. All fifteen first-year students and five of the seven second-year students scored below the 1,000-word band. Five upper-division students scored below 1,000; six scored between 1,000 and 2,000; eight scored between 3,000 and 4,000; and seven scored at the 5,000 level—i.e., close to 60% of the upper-division students had vocabulary levels of at least 3,000 words.

To determine the internal consistency of the RVLT and to provide an overall reliability estimate, Cronbach’s alpha was computed with the individual band scores as input. Cronbach’s alpha was 0.951, which indicates that the test’s internal consistency and reliability were very high. Table 6 shows Cronbach’s alpha hardly changed when bands were removed individually. This indicates that the test’s high internal consistency applied to all bands equally.

Separate paired-samples t tests were conducted to see if the differences between the individual band scores were statistically significant to provide support for the assumption that the five bands formed an implicational scale, i.e., that band scores got increasingly lower as the bands moved up from 1,000 to 5,000 words. Table 7 shows the item statistics for the individual bands, and Table 8 shows the results of the t tests.

Tables 7 and 8 show that there were statistically significant differences between the 1,000-and 2,000-word bands (1,000 mean = 19.56, 2,000 mean = 16.19, t = 3.92, p = 0.00, df = 47, d = 0.344), between the 2,000-word and 3,000-word bands (2,000 mean = 16.19, 3,000 mean = 12.98, t = 4.85, p = 0.00, df = 47, d = 0.30), and between the 4,000-word and 5,000-word bands (4,000 mean = 12.04, 5,000 mean = 9.48, t = 3.02, p = 0.004, df = 47, d = 0.253), but not between the 3,000-word and 4,000-word bands (3,000 mean = 12.98, 4,000 mean 508 FALL 2017

<table>
<thead>
<tr>
<th>Table 6: Cronbach’s Alpha, If Items Are Removed Individually</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale Mean If Item Deleted</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>1,000-word score</td>
</tr>
<tr>
<td>2,000-word score</td>
</tr>
<tr>
<td>3,000-word score</td>
</tr>
<tr>
<td>4,000-word score</td>
</tr>
<tr>
<td>5,000-word score</td>
</tr>
</tbody>
</table>
In summary, the results of the present study indicate that the 80% correct criterion is the best indicator of receptive mastery of a particular band for this form of the RVLT. They also show that the RVLT is a reliable test of receptive written vocabulary knowledge, even though it may not always be able to distinguish effectively between adjacent bands. Therefore, in the following reported results, the composite scores are primarily used.

**Reading Proficiency and Vocabulary Levels**

In the previous section, it was established that the scores of all five bands contributed equally well to the internal consistency and reliability of the RVLT. To establish how well the RVLT scores predicted the ACTFL reading proficiency levels, the composite vocabulary score consisting of the summed individual band scores was used.

A linear regression analysis was conducted to predict the ACTFL ratings from the vocabulary score. The assumptions of a regression analysis include a normal distribution of the data, an absence of outliers, and the homogeneity of variances between the two data sets. Figure 1 shows the P-P plot of the standardized residuals, examining the assumption of a normal distribution of the RPT and RVLT scores. While there is some evidence of nonnormality, the P-P plot by and large shows a linear relationship between the reading proficiency levels and the vocabulary scores. The maximum Cook’s Distance was 0.180 and the maximum Mahalanobis Distance was 2.72, supporting the assumption that there were no outliers. Figure 2 shows the scatterplot

<table>
<thead>
<tr>
<th>Bands</th>
<th>Mean</th>
<th>SD</th>
<th>SE Mean</th>
<th>Lower</th>
<th>Upper</th>
<th>t</th>
<th>df</th>
<th>p</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,000–2,000</td>
<td>3.38</td>
<td>5.96</td>
<td>0.86</td>
<td>1.64</td>
<td>5.11</td>
<td>3.92</td>
<td>47</td>
<td>0</td>
<td>0.344</td>
</tr>
<tr>
<td>2,000–3,000</td>
<td>3.21</td>
<td>4.59</td>
<td>0.66</td>
<td>1.88</td>
<td>4.45</td>
<td>4.85</td>
<td>47</td>
<td>0</td>
<td>0.300</td>
</tr>
<tr>
<td>3,000–4,000</td>
<td>0.94</td>
<td>3.82</td>
<td>0.55</td>
<td>0.55</td>
<td>2.05</td>
<td>1.70</td>
<td>47</td>
<td>0.095</td>
<td>0.087</td>
</tr>
<tr>
<td>4,000–5,000</td>
<td>2.56</td>
<td>5.88</td>
<td>0.85</td>
<td>0.85</td>
<td>4.27</td>
<td>3.02</td>
<td>47</td>
<td>0.004</td>
<td>0.253</td>
</tr>
</tbody>
</table>
between the studentized residuals and the predicted value of the standardized residuals. It shows a cloud of data scattered randomly and supports the assumption of homogeneity of variances.

After establishing that the assumptions for the regression analysis were met, we conducted the regression analysis to predict the ACTFL ratings from the vocabulary score. Pearson’s correlation between vocabulary score and reading proficiency was 0.915 with \( p < 0.001 \). The model explained 83.7\% of the reading results (\( R^2 = 0.837 \)). The linear regression analysis with reading proficiency as the dependent variable thus yielded a significant and large predictive effect of the vocabulary score on the reading proficiency rating: \( p < .001 \), intercept (\( \alpha \)): 0.25, slope (\( \beta \)): 0.06. Figure 3 plots the vocabulary (composite) scores and the reading proficiency levels expressed numerically and includes the results of the regression analysis. The regression analysis shows that the results of the vocabulary test predicted reading proficiency remarkably well; they explained 83.7\% of the variance, which is even higher than the highest variances for English established in previous research (62\% and 72\%; see the Vocabulary Size and L2 Development section above).

The extent to which particular vocabulary levels predicted reading proficiency levels was also examined. Table 9 shows the cross-tabulation of vocabulary levels and ACTFL reading proficiency levels. The table demonstrates that receptive mastery of the most frequent 1,000 and 2,000 words seems to predict a reading proficiency of Intermediate, while mastery of the most frequent 3,000 and 4,000 words seems to indicate a reading proficiency of Advanced. Finally, receptive mastery of the most frequent 5,000 words seems to indicate a reading proficiency of Superior. This is supported
by the equation of the regression analysis predicting reading proficiency levels on the basis of vocabulary scores rather than vocabulary levels. Table 10 shows the mean vocabulary scores associated with a particular level. These are the composite vocabulary scores that students at a particular vocabulary level received. In addition, Table 10 shows the results of the regression equation as a numeral and the ACTFL reading proficiency level associated with that numeral.

The results of the regression analysis in Table 10 provide further evidence that the 1,000-word and 2,000-word levels are associated with Intermediate; the 3,000-word and 4,000-word levels with Advanced; and the 5,000-word level with Superior (or Advanced High [AH]), suggesting that the Russian vocabulary levels may make even finer distinctions—i.e., the 1,000-word band predicting IL; the 2,000-word band predicting IM; the 3,000-word band predicting AL; the 4,000-word band predicting AM; and the 5,000-word band predicting at least AH.

**Discussion**

The reported data indicate that the RVLT exhibits high internal consistency and reliably measures a test taker’s breadth of vocabulary up to the 5,000 most frequent words of Russian. A correlation analysis between the composite vocabulary score and the vocabulary level established the 80% mastery criterion \( r = 0.96, p < 0.01 \) to be the better of the two criteria established in the literature for the RVLT. As Xing and Fulcher (2007) showed, two vocabulary-level tests created according to the same guidelines do not necessarily exhibit precisely the same level of difficulty. One reason presumably has to do with the fact that target words are randomly selected from the 1,000 words that constitute a band and may be biased toward the more frequent half of the band.
for one test rather than for the other. Therefore, the mastery criterion should probably be established empirically for each test separately. A correlation coefficient of 0.9 or higher certainly seems to be high enough. However, lower ones may also be acceptable.

The RVLT predicted 84.6% of the variance of students’ reading proficiency. This is one of the highest percentages

<table>
<thead>
<tr>
<th>Vocabulary Level</th>
<th>Less than 1,000</th>
<th>1,000</th>
<th>2,000</th>
<th>3,000</th>
<th>4,000</th>
<th>5,000</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Proficiency</td>
<td>NL</td>
<td>NM</td>
<td>NH</td>
<td>IL</td>
<td>IM</td>
<td>IH</td>
<td>AL</td>
</tr>
<tr>
<td>------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Superior</td>
<td>13</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>
reported in the literature and provides additional evidence of the reliability of the RVLT. The results of a regression analysis predicting the ACTFL reading proficiency levels on the basis of the overall vocabulary score consisting of all correct items of all five bands indicate that IL may be associated with knowing the most frequent 1,000 words, IM with the most frequent 2,000 words, AL with the most frequent 3,000 words, AM with the most frequent 4,000 words, and AH (and Superior) with the most frequent 5,000 words.

These correspondences track well with the lower end of the TORFL scale where the lexical minimum for Elementary (ACTFL Novice) is 780 words, Basic (ACTFL IM) is 1,300 words, and Level 1 (ACTFL IH) is 2,300 words. Beyond that, these results diverge from the TORFL’s stated lexical minimums. For example, at TORFL Level 2 (ACTFL AM), the stated lexical minimum for receptive vocabulary is 6,000 words. Some participants in this study scored at AM with a receptive vocabulary of 4,000 words and scored at Superior (TORFL Level 4) with receptive vocabulary scores showing knowledge of the 5,000 most frequent words. It is also worth noting, however, that since the RVLT only tests up to the 5,000-word frequency band, for participants who showed knowledge of this band it is impossible to know how many words beyond this they may also know. It is possible that the participants in this study who were rated Superior in reading proficiency and had 5,000-word vocabulary scores may well know many more than the 5,000 most frequent words.

The robust correlations between vocabulary knowledge and level of reading proficiency evidenced in these data are particularly informative when viewed in terms of students’ year of study and additional language learning experience. Among these test takers, no first-year student had mastered the 1,000-word level. In terms of reading proficiency, eleven first-year students were rated Novice, one was rated IL, and one was rated IM. Five of the seven second-year students had also not yet mastered the 1,000-word level. Four of these five students were rated Novice and one was rated IL. The two second-year students who had mastered the 1,000-word level had reading proficiency scores of IL and Intermediate High (IH). Among upper-division students, 75% of those who returned from an extended experience abroad and entered the third-year course sequence had mastered the 2,000-word level; 62.5% had mastered the 3,000-word level; 50% had mastered the 4,000-word level; and 29% had mastered the 5,000-word level. In addition, 58% were rated at least AL in reading and 29% were rated Superior; 71% of the Superior readers had a vocabulary level of at least 5,000, and the remainder had a vocabulary level of 4,000. The two upper-division students who were tested but who had not completed an extended experience abroad had strikingly different vocabulary and reading proficiency levels. One of them was rated IM with a vocabulary level of below 1,000

<table>
<thead>
<tr>
<th>TABLE 10</th>
<th>Results of Regression Equation Predicting ACTFL Levels on Basis of Vocabulary Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary Level</td>
<td>1,000</td>
</tr>
<tr>
<td>Mean Vocabulary Score</td>
<td>59</td>
</tr>
<tr>
<td>Result of Regression Equation</td>
<td>3.79</td>
</tr>
<tr>
<td>ACTFL Proficiency Level</td>
<td>IL</td>
</tr>
</tbody>
</table>
words. The other one was rated IH with a vocabulary level of 1,000 words.

From the perspective of L2 development, these findings are largely consistent with previous research showing that traditional classroom instruction alone is not sufficient to move Russian language students to Advanced levels of proficiency. Rifkin (2005, p. 11) found that “students without immersion experiences typically completed 4th-year Russian classes with only...intermediate-mid reading and speaking proficiency.” Davidson (2010) presented data on reading and listening proficiency gains for students who had completed Russian study abroad programs of varying lengths (2-month, 4-month, and 9-month programs). Overall, results showed that students began their programs with an IH reading proficiency and had moved into the Advanced range (summer and semester students) and Superior range (academic-year students) at the conclusion of their program. Tschirner (2016), reporting data from 152 Russian students at five institutions, showed that relatively low reading proficiency levels in Russian are not uncommon for U.S. college students. The mean second-year rating was NH (3.13) and the mean third-year rating was IM (5.09; Tschirner, 2016, p. 214).

Participant profiles in this study align with both types of students in these earlier studies; that is, with previous data on students whose language learning took place primarily in the domestic foreign language college classroom and on students who had extensive immersion experience. As in these studies, these results indicate that reaching Advanced-level reading proficiency requires input beyond regular classroom instruction. If Russian programs aspire to the goal cited in the Teagle Report (Modern Language Association, 2009)—namely, that Advanced-level reading proficiency should be the expected outcome for all language majors—then these findings about the role that vocabulary knowledge plays in reading proficiency are informative.

Most second-year students in this study were rated NH or IL. The two upper-division students who had not had an extended immersion experience had reading proficiency levels of IM and IH. If students are to reach the Advanced level by graduation, then better gains in reading proficiency are needed as students move through the curriculum. This study provides clear evidence that receptive vocabulary knowledge plays a critical role in a student’s overall level of reading proficiency, which in turn suggests the need to focus on vocabulary development. If indeed the 2,000-word level is a prerequisite for moving past the Intermediate reading proficiency level, then mastering receptively the most frequent 2,000 words of Russian can be considered an important element of any curricular reform. A focus on the development of receptive vocabulary for L2 Russian, or indeed any L2, may entail a combination of direct and indirect approaches to vocabulary learning as well as a reappraisal of what kinds of texts are made available to students at different points in the curriculum. Each of these issues is considered below.

First, these results suggest the need for programs and instructors to reconsider the place of vocabulary learning as well as the necessity of a more intentional approach to it. To what extent, for example, do the lexicons of the most commonly used textbooks build knowledge of words from the appropriate frequency bands and in the right amounts? How do textbooks and instructors frame the work of learning vocabulary? Are there features specific to a particular L2 that should be exploited to facilitate vocabulary learning? In answer to the first question, the lexicon of beginning- and intermediate-level textbooks may naturally gravitate to the language’s most frequent vocabulary items, but ideally this will be an explicit goal. On a related note, there may or may not be an expressed lexical goal for the level of study. If, as we have shown, to move
past Intermediate reading levels Russian learners need to have mastered at least the 2,000 most frequent words of the language, then pedagogical materials should explicitly target this mastery. The same would be true for materials in other languages. We may also reconsider which strategies are the most effective in promoting vocabulary learning and draw on the latest research to inform curriculum and pedagogy. For example, a recent study comparing vocabulary learning via rote memorization on the one hand and semantic mapping on the other showed “no significant difference ... between the vocabulary mean scores of the two groups on the posttest at the end of the four-month treatment period” (Khoii & Sharififar, 2013, p. 206). We are certainly not advocating a full-scale return to the memorization of decontextualized vocabulary lists, but perhaps a reassessment of the place of memorization as a vocabulary learning strategy is warranted. Finally, in the case of Russian in particular, pedagogical materials could capitalize more explicitly and systematically than they currently do on the language’s morphological structure to promote word learning. Russian words are built from roots, prefixes, and suffixes in a relatively transparent way. Making this evident to learners from the earliest stages can give them the ability to make predictions about new lexical items based on the component parts as they advance in their study. Many textbooks do a little of this. A good example of this approach can be seen in the Web-based introductory Russian textbook *Mezhdunamii* (http://www.mezhdunami.org) when it introduces the new lexical item кофеварка (kofevarka) ‘coffee maker’ with this explanatory note: “For example, кофеварка is made up of three parts: the familiar кофе [boil] + ва [boil] + ка = the thing for boiling coffee = coffee maker.” The authors point out to students that “[t]his might help you in learning some of the new words to see their components even though they are from unfamiliar words,” and they repeat this approach in multiple places. Empirical research that investigates the efficacy of this and other approaches is called for.

The second issue concerns the kinds of texts students are asked to read as they move through an undergraduate language program. The extensive literature addressing the relative merits of authentic vs. modified texts for L2 learners (e.g., Crossley, Louwense, McCarthy, & McNamara, 2007; Swaffar, 1985; Young, 1999) necessarily informs curricular choices concerning reading material. The results of the present study provide an additional element to consider, especially when they are viewed in light of the work cited earlier showing that a reader must know 95–98% of a text’s words in order to understand the text (Carver, 1994; Schmitt et al., 2011). Authentic literary texts are thus mainly beyond the reach of all but the most advanced students. Of course, reading itself can be a vehicle for vocabulary acquisition. However, an additional consideration when thinking about vocabulary acquisition via reading is research showing that approximately 12 repetitions of a word may be needed for the word to be acquired (Nation, 2014) and that to achieve 12 repetitions of the fourth most frequent 1,000 words in English, approximately half a million running words need to be read. Addressing these issues used to be the domain of the graded reader, which adapted and simplified authentic texts for lower-level readers. More recently, unsimplified authentic texts that have been augmented with scaffolding techniques designed to bridge the gap between known and unknown vocabulary are increasingly common (e.g., Comer, 2008). Both types of material (as well as carefully selected level-appropriate authentic texts) can provide the necessary structure and support that students require to build the receptive vocabulary that has been shown to be essential to advancing their level of reading proficiency.

Finally, it is important to note the limitations to the present study. First, because
the study relied on testing students who were currently enrolled in undergraduate language courses and because a goal of the research was to expand this area of inquiry to a less commonly taught language, the sample size (N = 48) is not large. However, the results were robust. Future research should seek to replicate and extend this work both to other students of Russian and to students of other languages. Second, while these results showed that students with extended stays in a Russian-speaking country tended to achieve Advanced or even Superior reading proficiency levels, not all students with similar amounts of experience abroad achieved such high scores. It was also the case that some students managed to reach a reading proficiency level with lower vocabulary scores than did their peers. Future research could look in more detail at individual student profiles to tease out possible causes of these two observed variations among students. For example, for the group who had spent a significant time in a Russian-speaking environment, it would be important to determine if the more successful readers actually spent more time reading or perhaps spent more time listening to Russian media and interacting in Russian. A comprehensive background questionnaire would provide valuable information to further contextualize these results.

Conclusion
This study investigated two sets of research questions: (1) What is the relationship between reading proficiency and vocabulary breadth in Russian? and (2) What levels of reading proficiency and vocabulary breadth may be reached after one, two, and three years of college Russian or after an extended immersion experience abroad? The present study found a significant and strong correlation between vocabulary size and reading proficiency: Mastery of the 1,000- and 2,000-word bands containing the most frequent lexical items of Russian was associated with the ACTFL Intermediate level of reading proficiency and the 3,000- and 4,000-word bands with the Advanced level; the 5,000-word band was associated with the Superior level. However, the higher levels of reading proficiency and vocabulary breadth were only reached by students returning from an extended immersion experience abroad (a missionary experience of 18–24 months). Students learning Russian only in college did not reach Advanced levels of reading proficiency. They averaged NM at the end of the first year and NH moving toward IL at the end of the second year. The two upper-division students who had not had an immersion period abroad were rated IM and IH. It was suggested that a major goal of any curricular reform should include an intentional approach to building receptive mastery of the most frequent 2,000–3,000 words in the first two college years through a mixture of direct vocabulary learning and practice and the use of graded readers and level-appropriate authentic texts, as has been suggested in the literature on vocabulary learning in the field of ESL.

Acknowledgments
This research was partly sponsored by the U.S. Department of Defense and the Institute of International Education. The publication does not necessarily reflect the positions or policies of the sponsoring organizations and does not constitute an official government endorsement.

References


Nation, I. S. P. (2014). How much input do you need to learn the most frequent 9,000 words? Reading in a Foreign Language, 26, 1–16.


Submitted April 25, 2017

Accepted June 4, 2017