AAUSC 2018 Volume—Issues in Language Program Direction

Understanding Vocabulary Learning and Teaching: Implications for Language Program Development

Johanna Watzinger-Tharp
University of Utah

Kate Paesani
University of Minnesota

Series Editors

Peter Ecke
University of Arizona

Susanne Rott
University of Illinois at Chicago

Editors

CENGAGE

Australia • Brazil • Mexico • Singapore • United Kingdom • United States

CENGAGE

AAUSC 2018 Volume: Understanding Vocabulary Learning and Teaching: Implications for Language Program Development

Peter Ecke, Susanne Rott, Johanna Watzinger-Tharp and Kate Paesani

Sr. Product Team Manager: Heather Bradley Cole

Product Assistant: Catherine Bradley

Marketing Manager: Sean Ketchem

Production Service: Lumina Datamatics, Inc.

Senior Designer: Sarah Cole

Manufacturing Planner: Fola Orekoya

For product information and technology assistance, contact us at **Cengage Customer & Sales Support, 1-800-354-9706** or **support.cengage.com.**

For permission to use material from this text or product, submit all requests online at **www.cengage.com/permissions.**

Library of Congress Control Number: 2018953450

Student Edition:
ISBN: 978-0-357-10668-6

Cengage
20 Channel Center Street
Boston, MA 02210
USA

Cengage is a leading provider of customized learning solutions with employees residing in nearly 40 different countries and sales in more than 125 countries around the world. Find your local representative at **www.cengage.com**

Cengage products are represented in Canada by Nelson Education, Ltd.

To learn more about Cengage platforms and services, visit **www.cengage.com.**

To register or access your online learning solution or purchase materials for your course, visit **www.cengage.com/student**

Printed in the United States of America
Print Number: 01 Print Year: 2018

Chapter 4

The Relationship Between Reading Proficiency and Vocabulary Size: An Empirical Investigation

Erwin Tschirner, Universität Leipzig
Jane Hacking, Fernando Rubio, University of Utah

Introduction

Reading proficiency is arguably the most important modality for academic achievement. Students need to be able to read and understand large quantities of academic texts such as introductory college textbooks, newspaper articles, scholarly articles, and reports of all kinds, including technical reports. Students of second languages (L2) also need to read literary texts such as novels, short stories, and poetry. Research on L2 reading is extensive and varied, addressing such questions as the role of L1 literacy (e.g., Bernhardt, 2010; Garrison-Fletcher, 2012; Sparks, Patton, Ganshow, & Humbach, 2012), grammatical knowledge (e.g., Berkemeyer, 1994; Koda, 1993), or vocabulary knowledge (e.g., Schmitt, 2008; Schmitt, Jiang, & Grabe, 2011) in the development of L2 reading proficiency. Bernhardt (2010) argues for a model of L2 reading that captures the contribution and interaction of many variables (e.g., L1 literacy, background knowledge, and processing strategies) to be able to move beyond partial explanations of the variance in L2 reading performance. Citing Brisbois (1995) Bernhardt notes, for example, that vocabulary knowledge accounts for 27% of this variance. We concur that a comprehensive model is important, but we also see the need for further understanding the role of specific variables such as vocabulary knowledge, especially in the L2 learning of languages other than English.

The role of vocabulary size as a predictor of reading proficiency has traditionally been conceived in terms of lexical frequency studies of texts, which examine the percentages of high- and low-frequency words the text contains. The emphasis on lexical frequency has at times led to daunting claims about how large one's vocabulary needs to be. For reading purposes, the figure most often cited is 8,000 to 9,000 word families consisting of approximately 24,000 individual lexical items, the knowledge which allows readers of English to read a broad range of texts (e.g., Nation, 2006). Hazenberg and Hulstijn (1996) estimated the minimum

vocabulary size needed to manage university study in Dutch as an L2 to be 10,000 base words (for a discussion of terminology, see below). They also found that the average vocabulary size of Dutch native speakers entering college in the Netherlands was 18,800 base words. Although figures of around 17,000 base words or word families are common assumptions regarding the vocabulary size of beginning college students (D'Anna, Zechmeister, & Hall, 1991; Goulden, Nation, & Read, 1990), a more recent study found that first-year British university students had an average vocabulary size of 10,000 word families (Milton & Treffers-Daller, 2013).

Vocabulary researchers use the terms "base words" (e.g., Hazenberg & Hulstijn, 1996) or "root forms" (e.g., Schmitt, 2010) to refer to the primary word forms to which inflectional and derivational affixes are added. The term "word family" includes the base/root form (e.g., work), its inflections (works, worked, working), and transparent derivations (worker). In addition, researchers try to measure two dimensions of vocabulary knowledge: *size/breadth* and *depth*. Measurements of vocabulary size typically seek to count the number of words a learner knows in a basic form–meaning way, whereas measurements of depth attempt to capture a learner's grasp of such things as a word's grammatical features, how it is used in sentences, its collocations, its associations, and constraints on its use.

Vocabulary size is clearly linked to the acquisition of competence in reading and to success in school learning in early childhood (Bornstein & Haynes, 1998; Ouellette, 2006). Milton and Treffers-Daller (2013) estimated that children learn approximately 500 new word families (in their L1) per year during their school years and that college students add about the same number of word families per year of study. They also noted that the resulting 10,000 base words at the beginning of college study still made it hard for these students to read introductory college textbooks and scholarly articles and suggested that student achievement may be explained as much by vocabulary size as by academic ability. Nation (2006) estimated that 8,000–9,000 word families are needed to reach the 98% text coverage of a wide range of nonacademic texts (novels, newspapers), which he suggested was necessary for reading efficiently and with adequate comprehension. In addition, Nation (2006) showed that the most frequent 1,000 word families in English average six types per word family. He therefore estimated that a vocabulary of 8,000 word families involves knowing approximately 34,000 individual word forms or types. For languages such as French or German, which have a much greater range of inflected word forms, a vocabulary of 6,500 base words may entail about 41,000 types (Kusseling & Lonsdale, 2013, p. 444). This finding is certainly relevant to the current study, which deals with three such highly inflected languages.

A major strand of both L1 and L2 vocabulary research concerns lexical thresholds, or how much vocabulary knowledge is required to achieve particular goals. For reading proficiency, this line of inquiry is commonly conceptualized in terms of text coverage, that is, how many individual words are typically included in certain kinds of texts such as novels and newspapers. Having established that a reader needs

to understand between 95% and 98% of the tokens of a particular text in order to understand it (Carver, 1994; Hu & Nation, 2000; Schmitt, 2008; Schmitt, Jiang, & Grabe, 2011), the question, then, was how many words one needs to know on average in order to understand the 95% or 98% of the tokens of similar texts. Corpus-based research showed, for example, that the most frequent 9,000 word families provided coverage of 98% of the tokens in a wide range of texts in English (Nation, 2006). The assumption, of course, is that a reader needs to understand 98% of the tokens to gain adequate understanding of a text. Schmitt, Schmitt, and Clapham (2001), however, did not find any threshold percentage, that is, a percentage at which comprehension dramatically improved, but rather a linear relationship between vocabulary knowledge and reading comprehension. Comprehension was measured using two instruments: a multiple-choice test "with an emphasis on items that required some inferencing skills in using information from the text" and a graphic organizer completion task, which "requires readers to recognize the organizational pattern of the text and see clear, logical relationships among already-filled-in information and the information sought through the blanks" (p. 31). Although text coverage of 95% was adequate for 60% comprehension and text coverage of 98% was adequate for 70% comprehension, they also found that text coverage of 90% still allowed a comprehension rate of 50% (Schmitt, et al., 2011, p. 35).

Another line of inquiry has attempted to correlate vocabulary size directly to reading proficiency by looking at the relationship between vocabulary breadth and the comprehension rate for a particular text. Reading proficiency and vocabulary size typically correlated strongly, between $r = 0.5$ and $r = 0.82$ (Laufer, 1992; Qian, 1999, 2002). Despite the fact that there were strong correlations between vocabulary size and reading comprehension, there was also a wide variation in the strengths of the correlations. On the one hand, this variation was probably due to the notion of "adequate comprehension" and the kinds of texts learners had to read, while on the other hand it may also have been due to the vocabulary size test used.

The two measures most commonly used in vocabulary size research are Nation's Vocabulary Levels Test (VLT) and Meara's XLex vocabulary test (Meara & Milton, 2003). Xing and Fulcher (2007) showed that two VLTs created using the same guidelines do not necessarily exhibit precisely the same level of difficulty. One reason for this discrepancy presumably has to do with the fact that target words are randomly selected from the thousand words that constitute a particular frequency band and may cluster more around the more frequent half of the band for one test than for the other.

There are several different mastery criteria used in vocabulary size studies, that is, the number of words the test taker needs to identify correctly in order to be considered to have mastered a particular band. The two most common percentages used are 80% (e.g., Xing & Fulcher, 2007) and 85% (e.g., Schmitt, Schmitt, & Clapham, 2001). Because test difficulty may vary, the mastery criteria should probably be empirically established for each test separately. We suggest

correlating the results of several mastery criteria (e.g., 75%, 80%, and 85%) with the overall raw score of the test and to select the mastery criteria that best correlates with the overall score.

To date, there is little empirical evidence relating vocabulary size to established frameworks of reference such as the Common European Framework of Reference for Languages (CEFR) or the American Council on the Teachers of Foreign Languages (ACTFL) Proficiency Guidelines. A handful of studies focusing on the CEFR suggested that the receptive knowledge of the most frequent 3,000 lexemes of a language is related to the CEFR B1 level, whereas knowledge of the most frequent 5,000 lexemes is related to the C1 level (Huhta, Alderson, Nieminen, & Ullakonoja, 2011; Meara & Milton, 2003; Milton, 2010).

Frameworks and Vocabulary Size

Although the ACTFL Reading Proficiency Guidelines (ACTFL, 2012) do not contain any explicit references to vocabulary size or content other than some very general descriptors such as "broad vocabulary" (Superior), "precise, often specialized vocabulary" (Superior), "high-frequency vocabulary" (Intermediate), "cognates and formulaic phrases" (Novice), there are published vocabulary lists, called vocabulary profiles, associated with the CEFR for many European languages. These lists are often staggering in size. For French, for example, the A1 level is associated with 975 base words, A2 with 1,645 words, B1 with 3,388 words, and B2 with 6,407 words (Kusseling & Lonsdale, 2013, p. 444). For Spanish, the Instituto Cervantes includes 1,146 words and phrasal words in its recommendations for A1, 2,730 words for A2, 6,066 words for B1, 11,830 words for B2, 14,910 words for C1, and 23,343 words for C2 (Hacking, Tschirner, & Rubio, in press). The lexical minimums for receptive purposes established by the Test of Russian as a Foreign Language are 2,300 for Level 1, which is thought to be equivalent to the CEFR level B1, 10,000 for Level 2 (= B2), 12,000 for Level 3 (= C1), and 20,000 for Level 4 (= C2) (Hacking & Tschirner, 2017). The *English Profile* established the following vocabulary sizes for the CEFR levels A1 to C2: 785 (A1), 2,382 (A2), 5,327 (B1), 9,502 (B2), 11,908 (C1), and 15,715 (C2) (Lahti, 2015). The Goethe Institute has published German word lists only for levels A1 to B1. The A1 list contains 650 words (Perlmann-Balme, n.d.) and the A2 list contains 1,300 words (Glaboniat, Perlmann-Balme, & Studer, 2016). The revised word list for the CEFR level B1 for German contains approximately 3,500 words (Glaboniat et al., 2016). For B2, the Goethe Institute refers to the word lists contained in *Profile Deutsch* (Glaboniat, Müller, Rusch, Schmitz, & Wertenschlag, 2005), which contains a total of 6,053 lexical items (excluding the multiword units that Profile Deutsch calls *Sprechhandlungen* [functions], which consist of phrases and short clauses). There are no suggestions for C1 and C2. Table 4.1 summarizes these assumptions for the languages discussed above.

Table 4.1. Vocabulary Size Assumptions for Various CEFR Levels by European National Test Institutes

	French	Spanish	Russian	English	German
A1	975	1,146	N/A	785	650
A2	1,645	2,730	N/A	2,382	1,300
B1	3,388	6,066	2,300	5,327	3,500
B2	6,407	11,830	10,000	9,502	6,053
C1	N/A	14,910	12,000	11,908	N/A
C2	N/A	23,343	20,000	15,715	N/A

These numbers are considerably different from language to language and appear to be very high at the higher levels, with B2 approaching or surpassing the 10,000 mark established by Hazenberg and Hulstijn (1996) for university study for several languages. The C2 level requirements for these languages, in fact, approach or surpass even native speaker's estimates of beginning college students. One reason for this wide disparity in vocabulary size assumptions may be the fact that the above lists were commonly not based on frequency studies but instead consisted of older lists, often from the 1970s, which were revised mostly on the basis of expert opinion. A comparison of the German B1 list with a frequency dictionary, for example, showed that the overlap between the two lists was only 60%—that is, 40% of the 3,500 words of the German B1 list did not belong to the most frequent 3,500 words of German, and 40% of the most frequent words of German were not included in the vocabulary list (Tschirner, 2017). Further, the CEFR vocabulary range descriptors focus mainly on production—predominantly oral production—in everyday language use, particularly at levels A1 to B2. For example, at the A2 level a learner "has a sufficient vocabulary for the expression of basic communicative needs," while at the B2 level he or she "has a good range of vocabulary for matters connected to his/her field and most general topics" and "can vary formulation to avoid frequent repetition, but lexical gaps can still cause hesitation and circumlocution" (Council of Europe, 2001, p. 112). This focus on production stems from the fact that the precursor of the CEFR, the Threshold Level project of the 1970s and 1980s, had a firm foundation in speech act theory, and it therefore conceived foreign language learning primarily as the acquisition of notions and (speech) functions. Tschirner (2017) argued that the CEFR approach focusing on production and speech functions is radically different from the receptive and frequency-based approaches common to vocabulary size studies in English as a second or foreign language (ESL/EFL). This difference affects how basic vocabularies are conceptualized, what words are considered important at what CEFR level, and how they are selected, and it produces vocabulary lists that are rather different from each other.

Research Questions

The focus of this inquiry is to establish vocabulary sizes required for various reading proficiency levels. The discrepancy between the results of the few existing empirical studies and the numbers established by corpus linguistic evidence is staggering. Because more and more researchers and curriculum specialists argue that vocabulary should be learned directly as well as indirectly, the question is how many words and what words. To provide evidence of the relationship between reading proficiency levels as defined by ACTFL and vocabulary size in German, Russian, and Spanish, the following research questions were addressed:

1. How well does reading proficiency as defined by ACTFL predict vocabulary size measured as the receptive knowledge of various bands of the most frequent 5,000 words in German, Russian, and Spanish?
2. What vocabulary sizes are predicted by what ACTFL reading proficiency levels in these languages?
3. What are the differences, if any, between German, Russian, and Spanish with respect to the relationship between reading proficiency level and vocabulary size?

Methods

Participants

Participants in this study were college students of German, Russian, and Spanish enrolled at two U.S. universities and at one university in Germany. There were 48 students of Russian and 52 students of Spanish enrolled at one of the two U.S. universities. There were 197 students of German. Ninety-seven of them were enrolled at the other U.S. institution, while one hundred students were enrolled at the German university. All U.S. students had American English as their first language (L1). Of the students in Germany, 75 spoke Arabic, 12 spoke Dutch, 11 spoke Thai, and two spoke Brazilian Portuguese as their L1. All participants took both the ACTFL Reading Proficiency Test (RPT) and the VLT. The Russian and Spanish tests were administered between September 2015 and January 2017, and the German tests were administered between June 2016 and June 2017.

Instruments

The ACTFL RPT is a standardized test for the global assessment of reading ability in a language (ACTFL, 2013). The test measures how well a person spontaneously reads texts when presented with texts and tasks as described in the 2012 ACTFL Proficiency Guidelines. The test formats used in this study consisted of 10 to 25 texts, depending on a participant's proficiency level. There were five sublevels: Intermediate Low (IL), Intermediate Mid (IM), Advanced Low (AL), Advanced

Mid (AM), and Superior (S). Each sublevel consisted of five texts accompanied by three tasks (items) with four multiple-choice responses, only one of which was correct. Test specifications included genre, content area, rhetorical organization, reader purpose, and vocabulary (ACTFL, 2013). Texts and tasks aligned at each level; for example, an Intermediate task required understanding information that was contained in one sentence, whereas Advanced tasks required the ability to understand information that was spread out over several sentences or paragraphs. Tasks and multiple-choice responses were in the target language and the test was web based.

The RPT is a timed test with a total test time of 25 minutes per sublevel. Two sublevels are scored together: either the two levels taken or, if more than two levels are taken, the two highest levels that can be scored according to the specific algorithm of the test. Because there are no Novice texts or tasks, the Novice levels are determined according to how close the test taker is to the Intermediate level. Test takers whose scores are below 33.3% of the maximum Intermediate score possible are rated NL, test takers whose score is between 33.3% and 50% are rated NM, and test takers whose scores are between 50% and 66.6% are rated NH. The test is Internet administered and computer scored (ACTFL, 2013).

The VLT consists of a receptive and a productive test (Institute for Test Research and Test Development, n.d.). It is modeled after the English VLT pioneered by Paul Nation (Nation, 1990). The VLT measures how many of the most frequent 5,000 words of German, Russian, and Spanish are known. It consists of five bands: the most frequent 1,000 words, 1,001 to 2,000 words, 2,001 to 3,000 words, 3,001 to 4,000 words, and 4,001 to 5,000 words. The receptive test, which was used in the present study, consisted of 10 clusters of six words each for each of these five bands. Each band was thus represented by 60 words. These words consisted of 30 nouns, 18 verbs, and 12 adjectives and were chosen at random from the 1,000 words of a band. Each cluster focused on one part of speech (e.g., nouns). Three words of a cluster were targets, which needed to be defined by choosing from a list of synonyms and paraphrases. The other three words were distractors. The maximum score per band was 30, that is, three points per cluster. Figure 4.1 provides an example of a cluster of the receptive VLT in English. Test takers are requested to select the word from column 2 that best matches the explanation in column 4.

1	business	A	part of a house
2	clock	B	animal with four legs
3	shoe	C	something used for writing
4	wall		
5	pencil		
6	horse		

Figure 4.1. Sample receptive Vocabulary Levels Test cluster

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). In addition to the mastery criteria of 80% and 85%, we added 75% to this study in order to determine which of the three percentages shows the strongest internal consistency and reliability of the vocabulary tests. Hacking and Tschirner (2017) argued that the mastery criteria should be established individually for each test, and they proposed to use the percentage that best correlates with the composite score consisting of the summed individual band scores. The maximum composite score for the five bands was 150, that is, five times 30.

Analysis

To examine the internal consistency and reliability of each VLT, Cronbach's alpha was computed with the individual band scores as input. If Cronbach's alpha was above 0.70, the VLT was considered to be internally consistent.

To determine the best mastery criteria for each VLT, that is, 75%, 80%, or 85%, the vocabulary levels of each student were calculated using these mastery criteria. First, raw scores were turned into percentages. These percentages were then used to provide a vocabulary level for each test taker. If a test taker, for example, scored 90% on the first band, 90% on the second band, 75% on the third band, 70% on the fourth band, and 40% on the fifth band, his or her vocabulary level was 3,000 if the mastery criterion was 75%, and it was 2,000 if the mastery criterion was 80% or 85%. The highest band that crossed the 75%, 80%, or 85% threshold was considered to be the respective vocabulary level of the test taker.

Spearman's rho was calculated in order to assess the relationship between students' composite vocabulary scores, consisting of all correct answers of all bands, and the vocabulary levels established for them by each mastery criteria. The mastery criteria correlating most highly with the composite score was used as the mastery criteria for that particular test.

To answer research question 1, separate linear regression analyses were conducted with ACTFL reading proficiency level as the predictor variable and vocabulary levels as the response variable for each language. The results of the regression equation for each ACTFL level were used to answer research questions 2 and 3.

Results

Reading Proficiency

The German and Spanish RPT were scored according to the ACTFL scale. The Russian RPT was originally scored according to the ILR scale and was rescored for this study using the algorithm used for ACTFL proficiency levels. Following Rifkin (2005) and others, ACTFL levels were coded numerically as follows: NL = 1, NM = 2, NH = 3, IL = 4, IM = 5, IH = 6, AL = 7, AM = 8, AH = 9, and S = 10. Figure 4.2 shows the distribution of the scores for the three languages.

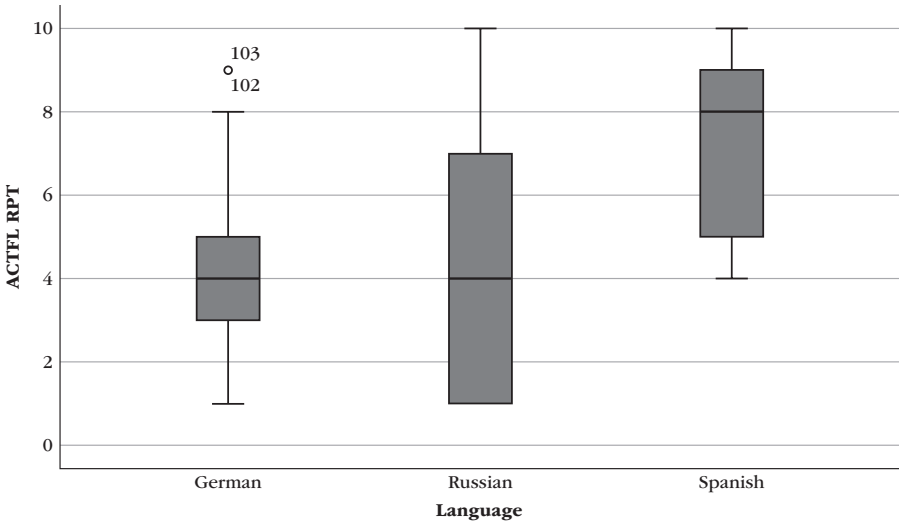

Figure 4.2. Distribution of ACTFL reading proficiency scores for German, Russian, and Spanish

Figure 4.2 shows that there were a few deviations from the normal distribution for each language. The distribution of German scores was leptokurtic, as can be seen by the interquartile range (box), which is relatively small: the central 50% of the scores varied from 3 (NH) to 5 (IM). In addition, there were two outliers at 9 (AH). The distribution of Russian scores was platykurtic. The central 50% of the scores varied from 1 (NL) to 7 (AL). In addition, it was positively skewed, that is, scores were bunched at the lower end of the scale (the NL level). Finally, the distribution of Spanish scores was negatively skewed; that is, the scores were bunched at the upper end of the scale (S). In general, however, the distributions are close enough to a normal distribution to make statistical inferences. The median (the line in the box) for German and Russian was similar (4 = IL), while it was considerably higher for Spanish (8 = AM), indicating that the Spanish students were, on average, much more proficient than the German and Russian students. The central 50% of the scores varied from 5 (IM) to 9 (AH). All three distributions, however, provided sufficient information, with data points ranging from 1 (NL) to 9 (AH) for German, 1 (NL) to 10 (S) for Russian, and 4 (IL) to 10 (S) for Spanish.

Vocabulary Size

Cronbach's alpha was computed with the individual band scores as input in order to determine the internal consistency of the three VLTs and to provide an overall reliability estimate. Cronbach's alpha is a measure of the internal consistency of the test, and it provides an estimate of the relationship between items and the interaction between subjects and items. In this case, each band was considered an item. Cronbach's alpha examines how closely related these bands are and whether

they could be considered to measure the same construct. In this sense, Cronbach's alpha may be deemed a measure of scale reliability. Cronbach's alpha levels above 0.70 are considered acceptable levels. Table 4.2 provides Cronbach's alpha for the three languages. Table 4.2 also provides the correlations between each mastery criteria—75%, 80%, and 85% correct—and the composite score consisting of the summed individual band scores in order to determine the mastery criterion that best correlates with the composite score.

Table 4.2. Cronbach's Alpha Coefficient of Reliability Computed between Bands ($p < 0.05$) and Pearson's r Correlations Between Composite Vocabulary Score of the Vocabulary Levels Test and Three Mastery Criteria: 75%, 80%, and 85%

	N	**Alpha**	**75%**	**80%**	**85%**
German	197	0.939*	0.905**	0.891**	0.853**
Russian	48	0.951*	0.960**	0.959**	0.923**
Spanish	52	0.951*	0.956**	0.956**	0.937**

* $p < .05$. ** $p < .01$.

Table 4.2 shows that Cronbach's alpha was statistically significant ($p < 0.05$) and very high for all three languages, indicating high internal consistency and reliability for the three vocabulary tests. Table 4.2 also shows that the mastery criterion best correlating with the composite score was 75% for German and Russian. For Spanish, the best correlation was tied between 75% and 80%. In the following, therefore, the mastery criteria of 75% will be used.

Reading Proficiency and Vocabulary Size

To provide an overview of the results, Tables 4.3–4.5 show cross tabulations of participants' ACTFL reading proficiency levels and vocabulary size for the three languages in question. Vocabulary sizes are indicated as follows: 0 = less than 1,000, 1 = 1,000, 2 = 2,000, 3 = 3,000, 4 = 4,000, and 5 = 5,000.

Table 4.3 shows that most test takers scored below the 1,000-vocabulary threshold for German. Most IL readers (N = 61, i.e., 79.2% of 77) scored below 1,000, while 18 (50% of 36) of the IM readers scored at 1,000 and above. All IH readers scored at 4,000 and above. It should be noted, however, that there were only three data points. All AL readers scored at 1,000 and above, and the majority of them (N = 8, i.e., 62% of 13) were at the 5,000 level. All AM and AH readers had a vocabulary size of at least 5,000.

Table 4.4 shows that approximately half of the test takers scored below the 1,000-vocabulary threshold in Russian. Half of the IL readers (3 of 6) scored below 1,000, the other half at 1,000. Fifty percent of the IM readers (3 of 6) scored above 1,000. AL and AM readers had vocabulary sizes of at least 3,000, and S readers had sizes of at least 4,000 words. Again, for some levels, there were very few data points.

Table 4.3. Cross Tabulation of Reading Proficiency Ratings and Vocabulary Levels for German

		RPT										
		NL	**NM**	**NH**	**IL**	**IM**	**IH**	**AL**	**AM**	**AH**	**S**	**Total**
Vocabulary Band	0	5	19	32	61	18	0	0	0	0	0	135
	1	0	0	2	12	2	0	2	0	0	0	18
	2	0	0	1	2	5	0	3	0	0	0	11
	3	0	0	0	2	4	0	0	0	0	0	6
	4	0	0	0	0	1	2	0	0	0	0	3
	5	0	0	0	0	6	1	8	2	2	0	24
Total		5	19	35	77	36	3	13	7	2	0	197

Note: NL = Novice Low, NM = Novice Mid, NH = Novice High, IL = Intermediate Low, IM = Intermediate Mid, IH = Intermediate High, AL = Advanced Low, AM = Advanced Mid, AH = Advanced High, S = Superior

Table 4.4. Cross Tabulation of Reading Proficiency Ratings and Vocabulary Levels for Russian

		RPT										
		NL	**NM**	**NH**	**IL**	**IM**	**IH**	**AL**	**AM**	**AH**	**S**	**Total**
Vocabulary	0	13	2	4	3	3	0	0	0	0	0	25
	1	0	0	0	3	0	2	0	0	0	0	5
	2	0	0	0	0	2	0	0	0	0	0	2
	3	0	0	0	0	1	0	2	1	0	0	4
	4	0	0	0	0	0	1	2	0	0	2	5
	5	0	0	0	0	0	0	1	1	0	5	7
Total		13	2	4	6	6	3	5	2	0	7	48

Table 4.5. Cross Tabulation of Reading Proficiency Ratings and Vocabulary Levels for Spanish

		RPT										
		NL	**NM**	**NH**	**IL**	**IM**	**IH**	**AL**	**AM**	**AH**	**S**	**Total**
Vocabulary	0	0	0	0	6	8	2	0	0	0	0	16
	2	0	0	0	1	0	0	0	0	0	0	1
	3	0	0	0	0	0	0	1	0	1	1	3
	4	0	0	0	0	0	0	5	3	5	3	16
	5	0	0	0	0	0	0	2	3	6	5	16
Total		0	0	0	7	8	2	8	6	12	9	52

Table 4.5 shows that most Spanish students (N = 35, i.e., 67.3% of 52) had vocabulary sizes of at least 3,000 words. Apart from one IL test taker, all Intermediate readers had vocabulary sizes of less than 1,000 words. Most of the Advanced readers had vocabulary sizes of at least 4,000 words, and most of the Superior readers had vocabulary sizes of at least 5,000 words.

In general, the Intermediate range seemed to be associated with a threshold vocabulary size of 1,000 words, while the Advanced range was associated with at least 3,000 or 4,000 words. The reason why German Advanced RPT levels seem to be associated with the 5,000 range may be due to the fact that most of the German students who had reached this vocabulary range were Dutch. The close relationship between Dutch and German may have given them an advantage with respect to vocabulary, but not necessarily with respect to reading. For Spanish, there seemed to be a gap between Intermediate and Advanced readers. Readers who were at the Advanced stage may have been able to take advantage of the vast shared vocabulary between English and Spanish, which may be particularly extensive at the 4,000 and 5,000 bands.

To determine the strength of the relationship between reading proficiency and vocabulary size and the predictive power of reading proficiency on vocabulary size, three separate linear regression analyses were conducted.

The simple linear regression analysis with reading proficiency as the predictor and vocabulary size as the response variable yielded a significant and large effect for all three languages. For German, it was as follows: $N = 197$, $R^2 = 0.531$, $p < 0.01$, intercept (α): –2.375, slope (β): 0.803. The model thus explained 53.1% of the vocabulary results, which is a large effect.

For Russian, the effect was also significant and even larger: $N = 48$, $R^2 = 0.796$, $p < 0.01$, intercept (α): –1.031, slope (β): 0.565. The model thus explained 79.6% of the vocabulary results.

For Spanish, the effect was also significant and also very large: $N = 52$, $R^2 = 0.736$, $p < 0.01$, intercept (α): –3.31, slope (β): 0.865. The model explained 73.6% of the vocabulary results.

To summarize for all three languages, the answer to research question 1 was that ACTFL reading proficiency strongly predicted vocabulary size, explaining between 53.1% (German), 73.4% (Spanish), and 79.6% (Russian) of the VLT results.

To answer the question of what vocabulary sizes are predicted by various reading proficiency levels (research question 2) and what differences there are between languages (research question 3), regression equations based on the above linear regression analyses were used to predict vocabulary sizes from ACTFL reading proficiency levels. For example, for Spanish the regression equation was $y = -1.031$ (intercept) $+ x * 0.565$ (slope), with y representing vocabulary size and x reading proficiency. Table 4.6 shows the vocabulary predictions by language and

ACTFL level. Because the ACTFL Novice level is characterized by understanding only isolated words and phrases, that is, it does not require any textual comprehension, the following table only includes ACTFL Intermediate, Advanced, and Superior levels.

Table 4.6. Vocabulary Size Predictions by ACTFL Reading Proficiency Levels

ACTFL Rating		IL	IM	IH	AL	AM	AH	S
Vocabulary Size	German	0.837	1.640	2.443	3.246	4.049	4.852	5.655
	Russian	1.229	1.794	2.359	2.924	3.489	4.054	4.619
	Spanish	0.114	0.970	1.826	2.682	3.538	4.394	5.250

Note: IL = Intermediate Low, IM = Intermediate Mid, IH = Intermediate High, AL = Advanced Low, AM = Advanced Mid, AH = Advanced High, S = Superior

Table 4.6 shows some clear patterns emerging for all three languages. Note that the figures provided represent the vocabulary ranges. For example, German IL = 0.837 indicates that IL predicts a vocabulary size of 83.7% of the 1,000 band, while IM = 1.640 indicates that IM predicts mastery of the 1,000 band plus 64% of the 2,000 band. Although there were some differences between the three languages at the IL and IM levels, IH seemed to be associated with the 2,000-word level, AL with the 3,000-word level, AM with the 3,000- and 4,000-word level, AH with the 4,000- and 5,000-word level, and S with the 5,000-word level. Although Russian started higher at IL and ended lower at S and Spanish started very low at IL, the general pattern of the regression analysis provides additional support for the postulation that the Intermediate level is associated with vocabulary sizes between 1,000 and 2,000 words and the Advanced level is associated with vocabulary sizes between 3,000 and 4,000 words.

Examining both the cross tabulation and regression results in response to RQ 2, the following seems to hold for all three languages: Intermediate Mid (IM) appears to be the threshold for the 1,000 band, while Advanced Low (AL) readers control the 3,000 band; Advanced Mid (AM) appears to be the threshold for the 4,000 band, while Superior (S) readers control, at least, the most frequent 5,000 words.

Discussion

Our study aimed at establishing the relationship between vocabulary knowledge and reading proficiency of L2 German, Russian, and Spanish learners. According to the 2012 ACTFL Proficiency Guidelines, readers with an ACTFL Advanced Mid (AM) reading proficiency are able to understand authentic narrative and descriptive texts such as "expanded descriptions of persons, places, and things and narrations about past, present, and future events." They understand the "main ideas,

facts, and many supporting details." They even "may derive some meaning from texts that are structurally and/or conceptually more complex." At the Advanced High (AH) level, they "begin to recognize author-intended inferences" and they have gained an "emerging awareness of the aesthetic properties of language and of its literary styles," which permits them to comprehend "a wide variety of texts" (ACTFL, 2012, p. 22). These descriptors seem to come close to what Schmitt (2008, p. 332) and others call "adequate comprehension" of "a wide variety of texts." Vocabulary researchers commonly associate this level of reading proficiency with a receptive vocabulary size of 8,000 to 9,000 words based on corpus linguistic studies. The few empirical studies relating vocabulary size with reading proficiency suggest that far fewer lexical items may be needed to gain adequate comprehension of a wide variety of texts. The results of the present study support these suggestions.

If we take the AM/AH levels to be indicative of the ability to achieve adequate comprehension of a wide variety of texts, then a vocabulary size of 4,000 to 5,000 words may be sufficient for German, Russian, and Spanish. Reading proficiencies of AM and AH predicted a knowledge of the most frequent 4,000 to 5,000 words in all three languages. It is important to note, however, that the German data are based on students with very different L1 backgrounds (Arabic and Thai as well as English and Dutch), which resulted in correlations that were much lower than the ones for Russian and Spanish. This may indicate that language distance is a variable that should be controlled in future studies.

Milton and Treffers-Daller (2013) suggested that children and young adults add approximately 500 base words to their L1 mental vocabulary per year in a largely intentional and explicit way. They question the explanations of vocabulary growth that primarily rely on implicit learning and argue that intentionally learning one to two new words per day is a thoroughly achievable task. Although L2 learners do not have the same amount of time as native speakers to grow their vocabularies, their task also seems a doable one. They probably need to learn 1,000 base words per year. And again, intentionally learning three to four new words per day appears to be a thoroughly achievable task. In addition, a more intentional approach to reading instruction that matches reading difficulty levels with the vocabulary size at a student's disposal may be required. In what follows, we examine some curricular implications suggested by both the results of our study and the literature review contained in this chapter.

Implications for Language Program Directors: An Intentional Approach toward Vocabulary Building and Reading Instruction

As mentioned in the introductory section, a common goal of language programs is the ability to read and analyze a variety of authentic texts, which typically requires an Advanced level of proficiency in the ACTFL scale. Tschirner (2016)

found that mean reading proficiency levels of college students were IH and AL in languages such as French, German, and Spanish in their fourth year of foreign language instruction, while the top 15% of students were AM, AH, and Superior. Reading may, in fact, be the only skill in which classroom L2 learners without extensive immersion experience may be able to reach the Advanced level. If, as we claim, there is a strong correlation between vocabulary knowledge and reading proficiency, it is imperative that language programs provide the necessary conditions to encourage and facilitate the acquisition of sufficient vocabulary to make the Advanced level an attainable goal of every language student. The most effective way to achieve that goal is to guarantee a clear articulation of vocabulary objectives across levels of instruction. Since this level of coordination is not common in most programs, language program directors (LPDs), in their position as overseers of the language curriculum, should play an active role in providing a well-articulated sequence of vocabulary learning objectives from beginning- to intermediate-, to advanced-level courses.

If we agree that reaching the 4,000 to 5,000 bands of vocabulary knowledge in the process of completing an undergraduate degree is a desirable goal, then approximately 1,000 words would need to be learned and retained per year. As Nation (2014) argues, a frequency approach will go a long way toward achieving the vocabulary sizes needed for effective reading proficiency—that is, the words studied in the first year should include the most frequent 1,000 of a language, the words studied in the second year should include the words from 1,001 to 2,000, and so on. The LPD, often the only expert in language acquisition in his/her program, can play a crucial role in helping faculty identify the vocabulary coverage needed to meet the objectives of each level in a language sequence and also in selecting appropriate materials that match the students' breadth of vocabulary. This process can begin with a basic assessment of students' vocabulary size, for example, by means of the VLT used in our study followed by an assessment of the appropriateness of the materials selected for each course. Cobb (2009) provides a number of suggestions that will help instructors become more aware of how vocabulary can determine the appropriateness of texts for learners and what vocabulary will be most useful to teach (see Cobb, 2009, for additional recommendations).

In addition to *what* vocabulary to teach, the LPD should also provide guidance and expertise on *how* to teach it. Research shows that becoming proficient at reading cannot be the result of exposure to input alone and that a combination of extensive reading and explicit vocabulary instruction may be the most effective recipe for success (Mondria, 2003; Schmitt, 2007). Although vocabulary depth—that is, the knowledge of how words are used, which includes knowledge about grammar, collocations, connotations, register, and style—still largely needs to be acquired experientially (i.e., while reading, writing, listening, and speaking), vocabulary breadth—that is, understanding the dictionary meaning of

words—may be learned by studying words directly using well-established activities that commit words to memory (Khoii & Sharififar, 2013). Grabe (2009) argues that in order to learn vocabulary, students need a combination of "vocabulary instruction, vocabulary-learning strategies, extensive reading and word learning from context, heightened student awareness of new words, and motivation to use and collect words" (p. 283). Nation (2013) maintains that a successful vocabulary learning program consists of 25% direct learning, 25% input-oriented learning (listening and reading), 25% output-oriented indirect learning (speaking and writing), and 25% fluency training. In other words, such a program requires the use of known vocabulary in differing contexts and under increased comprehension strain due, for example, to the time pressure of the communication situation. Nation also argues that graded readers rather than authentic materials help students acquire new vocabulary more effectively and efficiently, especially at the lower levels of language acquisition (Nation & Wang, 1999).

Although L2 teaching pedagogy has emphasized the importance of learning new words in context and doing extensive reading for vocabulary acquisition, the specific word-level learning strategies that, for example, Grabe (2009) mentions have been largely absent from most current, communicative-oriented language classrooms. The extensive body of research associated with vocabulary learning clearly suggests that direct vocabulary learning is not only beneficial but is also efficient, because it reduces the number of encounters needed with a new vocabulary item from 12 to perhaps three (Nation, 2014). The present study indicates that although the number of words one needs to know in order to read and understand most nonacademic texts, including literary texts, is still large, it is not as daunting as the corpus linguistic research seems to suggest. Rather than 8,000 to 9,000 words—as, for example, Nation (2014) postulates for reading newspapers and novels with sufficient ease and speed—4,000 to 5,000 words may be just as adequate, if the newspaper-and-novel level can be equated with ACTFL Advanced Mid or Advanced High.

As Nation and others have argued, studying words directly must be accompanied by activities in which words are used—in the present case, used in reading. Because it may be impossible to reach even 90% text coverage in reading passages with 1,000 or fewer words, an extensive reading approach based on controlled vocabulary materials, such as graded readers, would still be required. While there has been a renaissance in the production of graded readers in a number of languages over the past 10 years, these readers have usually not been developed based on word frequency studies but instead with a focus on short sentences, simple words, and simple syntax. However, there is a strong possibility that such texts will include high-frequency vocabulary. Highlighting these high-frequency items would provide students with the experiential lexical encounters they need in order to solidify their knowledge of the meaning of these words and begin their long-term acquisition of word depth. Graded readers are available in the three languages studied here.

The German Bookshop[1] provides a number of options for all levels and ages, some with accompanying audio materials. In Spanish, European Schoolbooks Limited[2] offers a variety of graded readers of different levels and genres, from comic books to literary classics. The same company also offers a number of options in Russian.[3]

When graded readers are not available or are inadequate for the learner's level of proficiency, we suggest taking advantage of technology in order to determine the degree of lexical complexity of a text and thereby its appropriateness based on the learners' vocabulary size. Cobb (2009) provides a complete overview of ways in which computational tools can facilitate vocabulary learning and the development of reading proficiency in general, mainly by matching learners' level of vocabulary knowledge to texts that are slightly above that level. The CALPER Corpus Portal developed by the Center for Advanced Language Proficiency Education and Research at Penn State University also provides a useful resource that describes in detail how to build a corpus and then use it for instructional purposes including vocabulary instruction.[4]

Matching texts to learners' vocabulary levels is an example of the type of lexical learning that Nation calls fluency training. Another approach toward achieving the text coverage needed for text comprehension and lexical learning is reading the same text several times, while focusing on different parts of its vocabulary load and focusing on it in various ways. While high-frequency vocabulary may be pretaught and be the explicit focus of vocabulary exercises associated with the reading passage, less frequent vocabulary necessary for understanding the passage and for achieving 90% text coverage may simply be glossed using a variety of techniques. Distinguishing between these two kinds of lexical items while reading, that is, words needed for text comprehension and words to be learned, allows learners to gain explicit knowledge of high frequency words in context while achieving the satisfaction of understanding a reading passage.

In sum, the guidance and expertise provided by the LPD are crucial for leading discussions about how the teaching and learning of vocabulary should progress between levels of instruction. Understanding the complex processes involved in L2 reading and the critical role of vocabulary knowledge in those processes is essential for providing an effective approach to vocabulary building and reading instruction.

Conclusion

Overwhelming evidence shows that there is a close relationship between the size of one's vocabulary and the ease with which different kinds of texts may be read.

[1] https://www.germanbookshop.co.uk/germanbooks/section/m4/c21
[2] https://www.eurobooks.co.uk/languagebooks/subject/SPA/m4/c21/6
[3] https://www.eurobooks.co.uk/languagebooks/subject/RUS/m4
[4] http://calper.la.psu.edu/content/corpus-portal

The question continues to be how large of a vocabulary learners need for various purposes and how it can be learned. While figures derived from corpus linguistic research have been rather daunting, consisting of 8,000 to 9,000 lexical units, the few existing empirical studies linking vocabulary size to reading proficiency seemed to question this conclusion. In addition, almost all studies looking at vocabulary sizes and what they allow people to do focused on English. The present study therefore took as its starting point languages other than English, in this case German, Russian, and Spanish, and it asked what vocabulary sizes are predicted by various reading proficiency levels. The present study provided additional strong evidence to support the argument that vocabulary size is a very strong predictor of reading proficiency, and it provided preliminary answers to what kinds of vocabulary sizes may be needed for various reading proficiency levels as measured by the ACTFL Proficiency Guidelines. For all three languages, the predictions were very similar, ranging from 1,000 to 2,000 words for the Intermediate levels to 3,000 to 4,000 words for the Advanced level and 5,000 words for the Superior level. We also argued that the best mix for quickly building up a sufficiently large vocabulary for reading purposes includes studying vocabulary directly while solidifying form–meaning comprehension and gaining vocabulary breadth knowledge through extensive reading (Day & Bamford, 1998). Among the most crucial aspects of a vocabulary program are the selection of the vocabulary and the selection of the reading texts used to solidify and expand the learning of words. The selection of vocabulary should be informed by frequency studies, while the selection of texts should be informed by the vocabulary load of the text, which calls for the use of graded readers at reading proficiency levels below the Advanced level.

References

ACTFL (2012). *ACTFL proficiency guidelines* [Electronic document]. Retrieved from http://www.actfl.org/sites/default/files/pdfs/ACTFLProficiencyGuidelines2012_FINAL.pdf

ACTFL (2013). *ACTFL reading proficiency test* (*RPT*): *Familiarization manual and ACTFL proficiency guidelines 2012—reading* [Electronic document]. Retrieved from https://www.languagetesting.com/pub/media/wysiwyg/ACTFL_FamManual_Reading_2015.pdf

Berkemeyer, V. (1994). Anaphoric resolution and text comprehension for readers of German. *Die Unterrichtspraxis*, *27*(2), 15–22.

Bernhardt, E. B. (2010). *Understanding advanced second-language reading*. New York, NY: Routledge.

Bornstein, M., & Haynes, O. M. (1998). Vocabulary competence in early childhood: Measurement, latent construct, and predictive validity. *Child Development*, *69*, 654–671.

Brisbois, J. E. (1995). Connections between first- and second-language reading. *Journal of Reading Behavior*, *27*, 565–584.

Carver, R. P. (1994). Percentage of unknown vocabulary words in text as a function of the relative difficulty of the text: Implications for instruction. *Journal of Reading Behavior*, *26*, 413–437.

Cobb, T. (2009). Necessary or nice: Computers in second language reading. In Z. Han & N. J. Anderson (Eds.), *Second language reading research and instruction: Crossing the boundaries* (pp. 144–172). Ann Arbor, MI: University of Michigan Press.

Council of Europe (2001). *Common framework of reference for languages*. Cambridge, UK: Cambridge University Press.

D'Anna, C. A., Zechmeister, E. B., & Hall, J. W. (1991). Toward a meaningful definition of vocabulary size. *Journal of Reading Behavior*, *23*, 109–122.

Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge, UK: Cambridge University Press.

Garrison-Fletcher, L. (2012). *The acquisition of L2 reading comprehension: The relative contribution of linguistic knowledge and existing reading ability* (Doctoral Dissertation). Retrieved from ProQuest Dissertations and Theses Database. (UMI No. 3499239).

Glaboniat, M., Müller, M., Rusch, P., Schmitz, H., & Wertenschlag, L. (2005). *Profile Deutsch*. Berlin, Germany: Langenscheidt.

Glaboniat, M., Perlmann-Balme, M., & Studer, T. (2016). *Goethe-Zertifikat B1: Deutschprüfung für Jugendliche und Erwachsene*, *Wortliste*. Munich, Germany: Goethe-Institut.

Goulden, R., Nation, I. S. P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, *11*, 341–363.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York, NY: Cambridge University Press.

Hacking, J. F., & Tschirner, E. (2017). The contribution of vocabulary knowledge to reading proficiency: The case of college Russian. *Foreign Language Annals*, *50*, 500–518.

Hacking, J. F., Tschirner, E., & Rubio, F. (in press). Vocabulary size, reading proficiency and curricular design: The case of college Chinese, Russian and Spanish. In S. Gass & P. Winke (Eds.), *Foreign language proficiency in higher education*. Cham, Switzerland: Springer.

Hazenberg, S., & Hulstijn, J. (1996). Defining a minimal receptive second language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics*, *17*, 145–163.

Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, *13*, 403–430.

Huhta, A., Alderson, J. C., Nieminen, L., & Ullakonoja, R. (2011, August). Diagnosing reading in L2: Predictors and vocabulary profiles. Paper presented at the ACTFL CEFR Conference, Provo, UT.

Institute for Test Research and Test Development (n.d.). *Vocabulary Tests*. Retrieved from http://www.itt-leipzig.de/static/startseite.html.

Khoii, R., & Sharififar, S. (2013). Memorization versus semantic mapping in L2 vocabulary acquisition. *ELT Journal*, *67*, 199–209.

Koda, K. (1993). Transferred L1 strategies and L2 syntactic structure in L2 sentence comprehension. *Modern Language Journal*, *77*, 490–500.

Kusseling, F., & Lonsdale, D. (2013). A corpus-based assessment of French CEFR lexical content. *Canadian Modern Language Review*, *69*, 436–461.

Lahti, L. (2015). Supplement to Lauri Lahti's conference article "Educational framework for adoption of vocabulary based on Wikipedia linkage and spaced learning." Retrieved from https://aaltodoc.aalto.fi/bitstream/handle/123456789/15382/J__lahti_lauri_2015.pdf?sequence=1&isAllowed=y

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126–132). London, UK: Macmillan.

Meara, P., & Milton, J. (2003). *The Swansea levels test*. Newbury, United Kingdom: Express.

Milton, J. (2010). The development of vocabulary breadth across the CEFR levels. In I. Bartning, M. Martin, & I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research*, *Eurosla Monographs* (Vol. 1) (pp. 211–232). N. P.: Eurosla. Retrieved from http://eurosla.org/monographs/EM01/EM01tot.pdf

Milton, J., & Treffers-Daller, J. (2013). Vocabulary size revisited: The link between vocabulary size and academic achievement. *Applied Linguistics Review*, *4*, 151–172.

Mondria, J.-A. (2003). The effects of inferring, verifying, and memorizing on the retention of L2 word meanings. *Studies in Second Language Acquisition*, *25*, 473–499.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York, NY: Newbury House.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, *63*, 59–82.

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge, UK: Cambridge University Press.

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*(2), 1–16.

Nation, I. S. P., & Wang, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, *12*, 355–380.

Ouellette, G. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *The Journal of Educational Psychology*, *98*, 554–566.

Perlmann-Balme, M. (n.d.). *Goethe-Zertifikat A1: Start Deutsch 1*, *Wortliste*. Retrieved from https://www.goethe.de/pro/relaunch/prf/et/A1_SD1_Wortliste_02.pdf

Rifkin, B. (2005). A ceiling effect in traditional classroom foreign language instruction: Data from Russian. *Modern Language Journal*, *89*, 3–18.

Schmitt, N. (2007). Current perspectives on vocabulary learning and teaching. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching*. New York, NY: Springer, 827–841.

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, *12*, 329–363.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. London, UK: Palgrave Macmillan.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the vocabulary levels test. *Language Testing*, *18*, 55–88.

Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, *47*, 484–503.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, *95*, 26–43.

Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2012). Do L1 reading achievement and L1 print exposure contribute to the prediction of L2 proficiency? *Language Learning*, *62*, 473–505.

Tschirner, E. (2016). Listening and reading proficiency levels of college students. *Foreign Language Annals*, *49*, 201–223.

Tschirner, E. (2017). Wortschatzwissen als Grundlage zweitsprachlicher Kompetenzen. In M. Clalüna & B. Tscharner (Eds.), *Bausteine des Spracherwerbs DaF/DaZ: Wortschatz—Chunks—Grammatik: Akten der Sechsten Gesamtschweizerischen Tagung für Deutschlehrerinnen und Deutschlehrer*, 17. und 18. Juni 2016 (pp. 11–21). Bern, Switzerland: AKDaF.

Xing, P., & Fulcher, G. (2007). Reliability assessment for two versions of vocabulary levels tests. *System*, *35*, 181–191.