

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



**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

© 2020, 2019 Cengage Learning, Inc.

Unless otherwise noted, all content is © Cengage

ALL RIGHTS RESERVED. No part of this work covered by the copyright herein may be reproduced or distributed in any form or by any means, except as permitted by U.S. copyright law, without the prior written permission of the copyright owner.

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

Chapter 5

Vocabulary Coverage and Lexical Characteristics in L2 Spanish Textbooks

Claudia Sánchez-Gutiérrez, University of California, Davis

Nausica Marcos Miguel, Denison University

Michael K. Olsen, Tennessee Technological University

Introduction

Textbooks are part of most educational settings and often guide the design of curricular content (Marcos Miguel, 2015; McDonough, Shaw, & Masuhara, 2013; McGrath, 2013). In the case of second language (L2) teaching, textbooks inform teachers' and Language Program Directors' (LPDs') choices regarding the vocabulary items to target in instruction (Allen, 2008). However, little is known about the characteristics of vocabulary in textbooks at the university level and how these characteristics might affect word learnability.

Previous literature has analyzed the extent to which textbooks follow a frequency criterion, as identified by frequency lists, when selecting vocabulary (Davies & Face, 2006; Lipinski, 2010). While this is a good starting point for research, it fails to account for lexical characteristics that might influence word learnability (Milton, 2009). For example, research shows that short words with concrete meanings are learned more quickly than long words with abstract meanings (Alsaif & Milton, 2012). Yet few studies have questioned the lexical characteristics of the words included in textbooks (e.g., Alsaif & Milton, 2012).

This study will include both a lexical frequency analysis and an investigation into the lexical characteristics of the words in the textbooks. Concretely, it will explore two lexical characteristics that affect word learnability, namely, word length and concreteness of meaning. Moreover, the evolution of word length and concreteness from elementary textbooks (ETs) to intermediate textbooks (ITs) will be investigated. This information is relevant for language teachers and LPDs, as it may help them select more learnable and useful lexical items for their curricula, make informed decisions on textbook adoptions, and better train their students to cope with lexical characteristics that increase the learning burden of a word (Laufer, 1990, 2012).

Lexical Frequency

While many vocabulary selection criteria for L2 classrooms could be proposed, lexical frequency has been one of the most used in the L2 acquisition literature (Horst, 2013; Nation, 2006; Schmitt & Schmitt, 2014). For example, Barcroft's (2012) input-based incremental vocabulary instruction approach suggests that vocabulary learning should not be reduced to incidental learning, but should rather be the result of a vocabulary plan based on a needs analysis or on frequency lists for the target language. For English, Barcroft points to the Academic Word List as well as more specific frequency lists for health or banking as useful resources for LPDs and other language teaching practitioners. Similar resources exist for other languages (see Appendix I).

Several authors have argued that learning the most frequent words in an L2 provides significant coverage of the linguistic input to which speakers of that language are exposed. Webb and Rodgers (2009a, 2009b) show that knowing the 3,000 most frequent words in English allows learners to understand over 95% of the words in TV shows and movies. Davies (2005) presents similar results in Spanish, noting that the 3,000 most frequent words in Spanish offer 94% coverage in oral texts. The level of coverage offered by these 3,000 words is smaller for written texts. For instance, Davies (2005) calculated that, in Spanish, the 3,000 most frequent words cover approximately 90% of words in written texts.

Schmitt, Jiang, and Grabe (2011) demonstrated that learners who know 95% to 98% of the words in a text are likely to understand it at 60%–68%.¹ Van Zeeland and Schmitt (2012) found that 95% of lexical coverage ensures a comprehension of about 75%. While knowledge of the 3,000 most frequent words in an L2 does not ensure complete comprehension of written and oral texts, compelling reasons nevertheless exist for prioritizing these words in the L2 classroom.

First, the frequency distribution of vocabulary in any given language is extremely skewed, with a few words being very frequent and covering most of the vocabulary. Meanwhile, a substantial number of words is rarely repeated, offering insignificant overall lexical coverage. As evidenced in Figure 5.1, the 1,000 most frequent words offer extremely broad coverage, while each of the next 1,000 words quickly becomes insignificant in terms of coverage. This explains why learning the first 1,000 words is so vital for L2 learners. However, a lexicon of merely 1,000 words is not sufficient to even approximate the coverage needed to facilitate text comprehension, and words in the next two bands of 1,000 words still offer an additional coverage of 3%–8% each. Frequency bands after the 3,000-word threshold, conversely, do not even reach 1% of coverage.

Additionally, Schmitt and Schmitt (2014) note that learner dictionaries of English generally comprise around 3,000 words, 90% of which are amongst the

¹This number was based on the results of a reading comprehension questionnaire completed immediately after the reading took place.

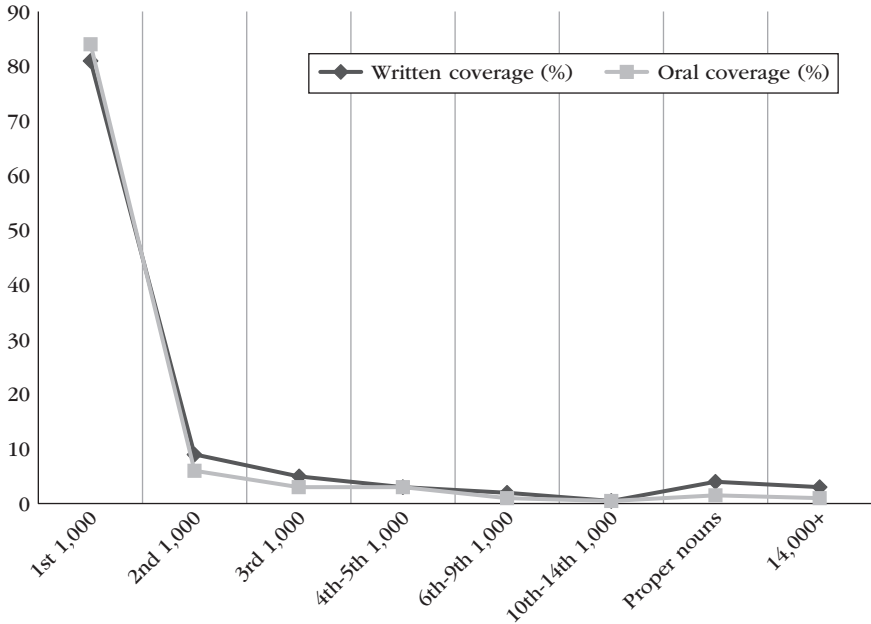


Figure 5.1. Percentages of text coverage by frequency (Nation, 2006, p. 79)

3,000 most frequent words. The authors argue that this finding demonstrates how word frequency and word usefulness, as judged by lexicographers, are not at odds, but actually tend to cohere. If the words chosen by lexicographers to be part of learner dictionaries correspond, for the most part, with the 3,000 most frequent words in the language, it seems obvious that those words should be given priority in the classroom. Finally, the authors show that most L2 English graded readers² contain around 3,000 different words, suggesting that the acquisition of 3,000 words is a reasonable goal for learners who wish to be equipped for reading non-adapted L2 texts. All these arguments indicate that teachers should prioritize the most frequent 3,000 words, as these words provide a solid lexical base for understanding most English L2 texts that students will encounter. These arguments might be applied to other L2s. For instance, Davies' (2005) study shows similar percentages of coverage at each frequency level for Spanish.

In the context of American universities, where language courses are generally taught for two years, after which students are expected to enroll in target-language literature, culture or linguistics classes, the 3,000 most frequent words should be taught during those first two years to the extent possible. Gairns and Redman (1986), reported by Milton (2009), "suggest an average of 8 to 12 productive items

² These are addressed mostly to elementary and intermediate learners, as more advanced learners are expected to read unmodified texts.

per class as representing reasonable input, which might lead to over 1000 items being presented in 125 hours of tuition” (p. 196). Therefore, in a semester of 14 weeks with three hours of instruction per week, a teacher could present 336–504 words per semester, 672–1,008 per year,³ and 1,344–2,016 in two years.

Thus, while time constraints may prevent students from learning all 3,000 words, the concept of frequency and the goal of 3,000 words can nevertheless serve as guiding principles for vocabulary selection. It is not a matter of teaching all 3,000 words, but of ensuring that most words taught belong to the list of most frequent words.

Impact of Lexical Characteristics in Learning and Processing Burden

Although frequency represents a useful measure for selecting vocabulary, it does not provide information on the learning burden of those words. Indeed, the most frequent words are not necessarily the easiest to learn and, conversely, infrequent words could potentially be extremely easy. For this reason, Laufer (1990) suggests that word learnability should also be taken into account when selecting vocabulary for the L2 classroom. She states that:

When words are easy to learn, they should be taught even if, on the basis of the frequency/range principles, they would not be considered useful. Cognates, words related structurally to already familiar words, and words with exact L1 equivalents all may require little learning effort and at the same time increase the communicative ability of the learner considerably. (p. 150).

LPDs, textbook authors, and language instructors should thus reflect on their criteria for selecting target vocabulary: frequency is important, but learner characteristics and word learnability must be considered as well. Of the many factors affecting word learnability, similarities between the L1 and the L2 (e.g., cognateness and L1 equivalency) may seem like the best complement to frequency when selecting target vocabulary. Yet this selection criterion cannot be the main or the only one. If we take English as the L1 by default, and thus assume that students will learn Spanish–English cognates faster than noncognates, we might favor L1 English students over learners with different linguistic backgrounds (see Szubko-Sitarek, 2011). Given the increasingly diverse population of students in American universities (Institute of International Education, 2015), a vocabulary list based only on frequency and L1/L2 similarities may disadvantage certain multilingual learners. Thus, when selecting vocabulary, it is important to also consider learnability characteristics that are not related to the learner’s L1.

³ These numbers would include instances of both explicit teaching and incidental learning, given that several words, as articles or classroom management words will be repeated class after class and might not require as much explicit attention.

Several authors have investigated such characteristics (Alsaif & Milton, 2012; Masrai & Milton, 2015; Willis & Ohashi, 2012) and have found an increased learning burden for words that are long, as opposed to short (Alsaif & Milton, 2012; Peters, 2016; Willis & Ohashi, 2012), and abstract, as opposed to concrete (Alsaif & Milton, 2012; De Groot & Keijzer, 2000; Ellis & Beaton, 1993). Thus, the longer and more abstract a word is, the more difficult it is to retain, whereas the shorter and more concrete it is, the easier it is to retain. The form *dog* will be more easily remembered than *institutionalization*.

However, the difficulty of long and abstract words can be tackled with two simple strategies. On the one hand, while words in Spanish tend to be longer than their English counterparts (Cantos & Sanchez, 2011), this is mostly due to the reliance of the Spanish language on suffixation as a means of creating new words (Lang, 2009). Thus, training the students in the recognition and use of productive affixes and word formation rules in Spanish could go a long way in addressing the difficulties that word length might represent (Morin, 2003, 2006). On the other hand, words with abstract meanings are more difficult to learn because they cannot be directly associated with images or sensorial experiences. Indeed, abstract vocabulary is so difficult to learn that a popular vocabulary-acquisition strategy, the keyword technique (see Hulstijn, 1997; Ecke, 1999, 2004 for a review), encourages learners to visualize a concrete object that they can link to an abstract concept. For example, a Spanish speaker learning the word *canny* in English could choose to associate it with the similar-sounding *can*, which is a Spanish word for *dog*. Then, he could elaborate the image of a dog with a pipe looking like a detective, which would be easy to connect to the meaning of *canny* as “clever.” In this way, the abstract word *canny* is now attached to an image that bridges in a more concrete way the newly encountered word in the L2 with the L1.

Textbook Analyses of Frequency, Concreteness, and Length

Word frequency in L2 English textbooks is relatively well researched (Alcaraz Marmol, 2009; Criado & Sánchez, 2009; Donzelli, 2007; Milton, 2009). Most of these studies determine the extent to which textbooks include vocabulary items from the first 2,000–3,000 most frequent English words and above.⁴ For example, Donzelli (2007) analyzed a textbook utilized in a primary school for Italian-speaking children who were learning English. The textbook included half of the items from the 1,000-word band, 17.80% of the items from the 2,000-word band, and 30% from the 2,500–3,000-word band. The presence of low-frequency words (over the

⁴The program Range (Nation, n.d.) is widely used for the purpose of determining what percentage of words in a text is drawn from which band of the word list. The New General Service List, the Academic Word List, and the British National Corpus (BNC) or the Corpus of Contemporary American English (COCA) are the main tools for measuring frequencies in English as an L2. For more information, visit: <https://www.victoria.ac.nz/lals/about/staff/paul-nation>

2,500 mark, in this case) led Donzelli to conclude that the book offered a rich lexical input where word frequency was not the leading criterion in vocabulary selection. Some examples of “unusual words” in the textbook were *basketball*, *soccer*, *homework*, *chicken*, and *geese*. Although these are not high frequency in the wordlists utilized, they are pertinent words for students that age, according to the author.

Similarly, Alcaraz Mármol (2009) found 50% of words from the 1,000-word band, 16% of words from the 2,000-word band, and 40% of words from the 2,000–3,000-word band in a primary textbook utilized for Spanish-speaking children learning English in Spain. She also separately analyzed each textbook chapter and found different percentages by chapter. Nevertheless, words in the 1,000 band were consistently the most frequent, representing between 53% and 80% of the words in each chapter. As was the case with Donzelli’s (2007) findings, the words above the 2,000-band referred to specific content targeting young learners, such as *hamster*, *schoolbag*, *spider*, *pumpkin*, or *zoo*.

Compared to English, there are fewer studies on frequency in Spanish L2 textbooks (Davies & Face, 2006; Godev, 2009; López Jiménez, 2014). Two previous studies analyzing frequencies in adult L2 Spanish textbooks are Davies and Face (2006) and Godev (2009).⁵ Davies and Face explored the vocabulary chapter lists (i.e., the list of explicit targets at the end of each chapter) of six Spanish textbooks: three first-year and three second-year textbooks used in American universities. These textbooks included 10%–50% of the 2,000 most frequent Spanish words. As was the case in other studies (e.g., Lipinski, 2010; see Milton, 2009), vocabulary selection varied by book. For example, a book such as “Dos Mundos” had 3,217 words, of which only 50% belonged to the 1,000–3,000 band, whereas only 28% of the 1,689 words in “Mundo 21” were in the 1–1,689 band. Thus, textbook authors seemingly disregard frequency when selecting target items and cover only a relatively small percentage of the most frequent Spanish words. These results were confirmed in Godev’s (2009) study, which also suggested that frequency played little to no role in the selection of vocabulary for five first-year college Spanish textbooks published in the United States.

This situation is not exclusive of Spanish and English L2 textbooks, as Lipinski (2010) also found that 50% to 60% of the words in the 1,000 band were included in three German L2 textbooks for first- and second-year programs. Words from the 2,000- and 3,000-word bands were represented only minimally in these textbooks, at approximately 30% and 15%–20%, respectively. These results are disheartening, given the importance of focusing on those 3,000 most frequent words in the L2 classroom (see Horst, 2013; Schmitt & Schmitt,

⁵ In most studies on Spanish L2, Davies’ (2006) *Frequency Dictionary* is used for frequency assessments of textbook vocabulary, as it presents the 5,000 most frequent words of the Spanish language, calculated from a version of the *Corpus del Español*, which contains 20 million words (Davies, 2002).

2014). Moreover, low-frequency words are common in L2 textbooks because vocabulary selection is generally based on semantic clusters (e.g., food, family, and free time) rather than frequency (Davies & Face, 2006; López Jiménez, 2014).

While the role of frequency in selecting textbook vocabulary items has been studied often, it is rarely studied in combination with the concreteness and length of target items. Indeed, to the best of our knowledge, only one such study has been conducted that pointed in that direction. Alsaif and Milton (2012) analyzed 22 English L2 textbooks, from beginner to advanced levels, used in Saudi Arabia primary and secondary schools. When the included vocabulary items were added together as a sum, all 22 textbooks covered “just over 80% of the 2000 most frequent words” and “half of the most frequent 5000 words” (p. 26) in seven years of English courses. Additionally, Alsaif and Milton noticed that little new vocabulary was added in higher proficiency textbooks that was not already present in lower proficiency books. This insufficient inclusion of appropriate vocabulary at increasing proficiency levels might explain the stagnating vocabulary level observed by Alsaif (2011) in students of English in Saudi Arabia. Finally, the authors reported that shorter and more concrete words in the textbooks were learned better by the students. Alsaif and Milton (2012) interpret this finding as an argument in favor of introducing not only high-frequency words in the textbooks but also words that are short and concrete, whenever these fit into a communicative context that is relevant for the course.

Objectives and Research Questions

This study aims to complement previous research in two ways. First, it will analyze word frequency, concreteness, and word length in a larger corpus of textbooks than any previous studies and will be the first to look at these three variables in Spanish L2 textbooks. Additionally, it will explore the extent to which words in ET and IT textbooks differ from one another in terms of these three characteristics. While Alsaif and Milton (2012) have already looked at the percentage of words from different frequency bands added to textbooks as proficiency levels increase, no study to date has analyzed the evolution of word concreteness and length at different proficiency levels.

Accordingly, this study addresses the following research questions (RQs):

1. To what extent do elementary and intermediate Spanish textbooks cover the 3,000 most frequent words in Spanish?
2. Is there an increase in the number of less frequent words from elementary to intermediate textbooks?
3. Are words in intermediate Spanish textbooks more abstract and longer than words in elementary Spanish textbooks?

Methods

Textbooks

Sixteen textbooks used in American Universities were selected for this study. Eight textbooks were ETs, designed for students with novice proficiency levels, and eight were ITs, designed for students at novice-high and intermediate levels of proficiency on the ACTFL proficiency scale. ETs are generally used during the first year of university-level Spanish courses, whereas ITs are used during the second year of Spanish instruction. Table 5.1 lists the specific books included in each proficiency level.

Processing of the Textbook Glossaries

The glossary at the end of each of the 16 textbooks was scanned and saved in plain text format. Additionally, the list of the 20,000 most frequent words in Spanish was downloaded⁶ from the website of the *Corpus del Español* (Davies, 2002) and divided into four frequency bands: band 1 was composed of the first 1,000 most frequent words in Spanish, band 2 contained the next 1,000 most frequent words, band 3 contained the next 1,000 most frequent words, and the low-frequency band included all the words that were not among the 3,000 most frequent words in the corpus. AntWordProfiler (Anthony, 2014) was used to perform the frequency analyses: each glossary was entered in the software as a User File and the lists of

Table 5.1. Textbooks Analyzed in this Study

Elementary Textbooks	Intermediate Textbooks
<i>Adelante</i> , 2nd ed. (2015), Vista Higher Learning	<i>Anda</i> , 2nd ed. (2013), Pearson (<i>EC</i>)
<i>Arriba</i> , 6th ed. (2015), Pearson	(<i>AC</i>) <i>Atando Cabos</i> , 4th ed. (2012), Pearson
(<i>CB</i>) <i>Con Brío</i> , 3rd ed. (2013), Wiley	<i>Conexiones</i> , 5th ed. (2014), Pearson
(<i>DyH</i>) <i>Dicho y Hecho</i> , 10th ed. (2015), Wiley	<i>En Comunidad</i> (2008), McGraw-Hill
(<i>DM</i>) <i>Dos mundos</i> , 6th ed. (2006), McGraw-Hill	<i>Enfoques</i> , 4th ed. (2016), Vista Higher Learning
<i>Nexos</i> , 3rd ed. (2013), Cengage	<i>Fusión</i> (2010), Pearson
(<i>PV</i>) <i>Pura Vida</i> (2014), Wiley	<i>Imagina</i> , 3rd ed. (2015), Vista Higher Learning
<i>Vistas</i> , 4th ed. (2012), Vista Higher Learning	(<i>PyA</i>) <i>Punto y Aparte</i> , 5th ed. (2015), McGraw-Hill

⁶This is not a free service, but the list of the first 5,000 most frequent words in Spanish will soon be available in the second edition of the *Frequency Dictionary of Spanish* (Davies & Davies, 2018). The second edition will represent a significant improvement compared with the first one (Davies, 2006), as frequency data are obtained from a corpus of two billion words, extracted from websites and blogs from all over the Spanish-speaking world. Additionally, data were gathered during 2013–2014, which ensures that frequency counts are based on the current state of the language.

words from each frequency band were entered as Levels lists.⁷ Using these data, the program sorted the words from each glossary into the appropriate frequency list and tallied the number of words from each glossary belonging to each band.

Additionally, the concreteness and word length data were obtained from EsPal (Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013), an online repository of lexicometric information (e.g., length, frequency, concreteness, and familiarity) calculated from a corpus of over 700 million words in Spanish. Length was computed as the number of letters in a word, while concreteness data were obtained by asking native speakers of Spanish to rate the concreteness of words on a scale from 1 (extremely abstract) to 7 (extremely concrete).

Results

Lexical Frequency in Textbooks

Two aspects of the frequency of words in ETs and ITs were studied to answer RQ1: (1) the distribution of words in each textbook by frequency bands (i.e., 1, 2, 3, and low frequency) and (2) the percentage of words from each band in each textbook. These are complementary ways of looking at the data, as the former gives an insight into the coverage offered by words from each band in each textbook, while the latter indicates a distributional proportion of words from each band.

This distinction is relevant because the first measure is dependent on the length of the glossary. For example, if a textbook with a glossary of 4,500 words includes 900 of the 1,000 most frequent words, these will barely represent 20% of the total glossary. This result might give the false impression that the book includes few words from band 1. Conversely, if those same 900 words are part of a glossary that contains 1,200 words, the proportion will be 75%. Thus, while the level of coverage offered by words in a specific frequency band provides important information, such data should be analyzed in the light of a second result: the number of words at each frequency band included in the textbook. Ideally, a textbook that focuses on the 3,000 most frequent words in Spanish should follow these two premises: (1) words from bands 1 to 3 offer a high level of coverage in the glossary, and (2) most words from those frequency bands are included. ETs should include most of the words from band 1 and around 50% of those in band 2, while ITs should include the other 50% of words in band 2 and most words from band 3. Table 5.2 describes the number and percentage of coverage of words in each frequency band in ETs.

As can be observed in Table 5.2, words from bands 1 to 3 represent over 50% of the words in all ETs, with four books reaching a coverage of over 60% for those words: *Adelante*, *Con Brío*, *Dicho y Hecho*, and *Vistas*. However, Figure 5.2 shows

⁷ For more information on the program, see <http://www.laurenceanthony.net/software/antwordprofiler/>

Table 5.2. Total of Words and Distribution of Words by Frequency Band in Elementary Texts

	<i>Adelante</i>		<i>Arriba</i>		<i>CB</i>		<i>DyH</i>		<i>DM</i>		<i>Nexos</i>		<i>PV</i>		<i>Vistas</i>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
1	576	33.4	729	24.8	381	32.3	408	32.7	832	21.6	534	28.9	379	33.2	438	33.8
2	303	17.6	535	18.2	222	18.8	223	17.9	638	16.6	316	17.1	181	15.9	222	17.1
3	197	11.4	387	13.1	133	11.3	131	10.5	481	12.5	228	12.4	125	11	149	11.5
1-3	1,076	62.4	1,651	56.1	736	62.3	762	61	1,951	50.6	1,078	58.4	685	60	809	62.4
> 3	647	37.6	1,292	43.9	445	37.7	487	39	1,902	49.4	768	41.6	456	40	488	37.6
Total	1,723		2,943		1,181		1,249		3,853		1,846		1,141		1,297	

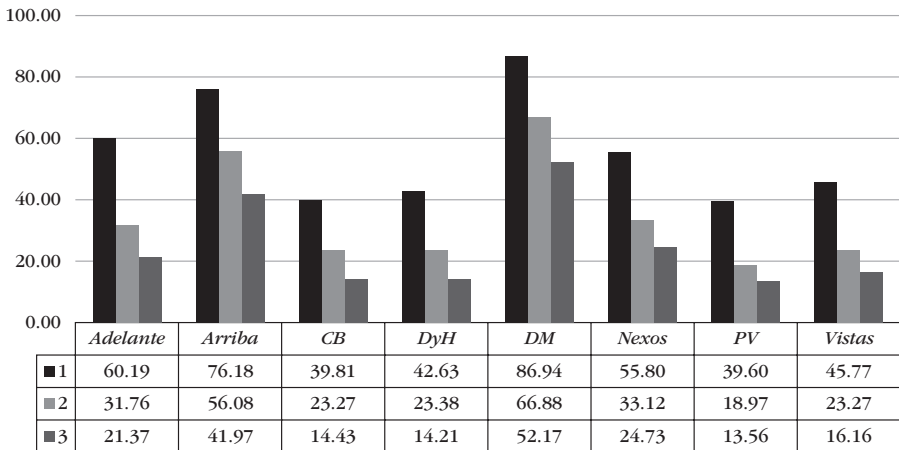


Figure 5.2. Percentages of words from frequency bands 1 to 3 included in elementary texts

that only *Adelante* includes more than 50% of the words in band 1. Thus, none of the books meet our two criteria: namely, they do not include most of the words in band 1 nor approximately half of those in band 2, and bands 1 to 3 do not offer a maximal coverage of the vocabulary in the textbooks.

Other textbooks, such as *Dos Mundos* or *Arriba*, include a broader selection of words from bands 1 to 3 than *Adelante*. However, in these cases, the presence of a high number of words from a specific frequency band probably does not result from a clear selection criterion but rather from the length of the glossaries. Therefore, more words from all bands, independent of frequency, are included.

Among the ITs, (Table 5.3) only *Anda* presents a coverage of over 60% for words in bands 1 to 3. However, due to its short glossary (1,007 words), the actual number of words included from band 1 is so low that it does not even cover 40% of the 1,000 most frequent words in Spanish, as can be observed in Figure 5.3. The situation for words in bands 2 and 3 is no better, with a proportion of 19.71% and 11.39%, respectively.

Table 5.3. Total of Words and Distribution of Words by Frequency Band in Intermediate Texts

	<i>Anda</i>		<i>AC</i>		<i>Conexiones</i>		<i>EC</i>		<i>Enfoques</i>		<i>Fusión</i>		<i>Imagina</i>		<i>PyA</i>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
1	349	34.7	350	29	350	21.6	833	22.6	434	23.1	237	19.2	629	27.6	870	20.5
2	188	18.7	195	16.2	261	16.1	620	16.8	322	17.2	185	15	410	18	693	16.3
3	105	10.4	143	11.9	205	12.7	468	12.7	248	13.2	128	10.4	288	12.6	575	13.6
1-3	642	63.8	688	57	816	50.4	1,921	52.1	1,004	53.5	550	44.5	1327	58.1	2,138	50.4
>3	365	36.2	518	43	802	49.6	1,768	47.9	872	46.5	685	55.5	956	41.9	2,104	49.6
Total	1,007		1,206		1,618		3,689		1,876		1,235		2,283		4,242	

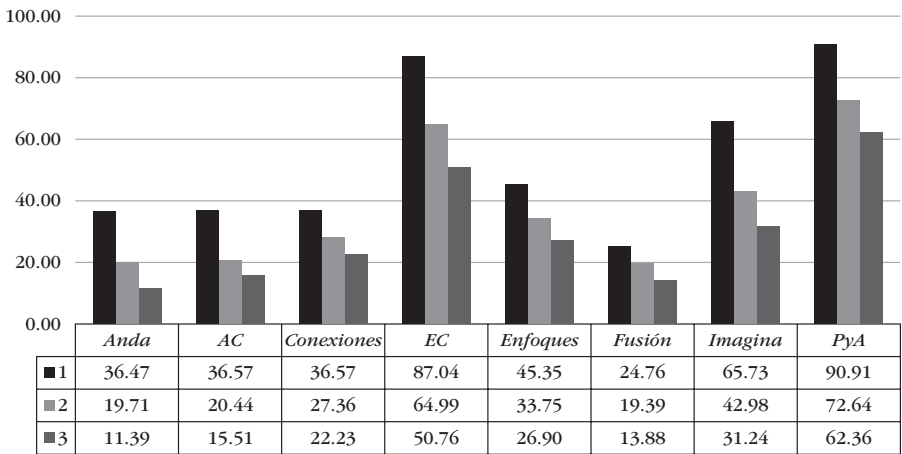


Figure 5.3. Percentages of words from frequency bands 1 to 3 included in intermediate texts

Thus, while this textbook seemingly offers good coverage, this perception was due only to the limited length of the glossary, not to a clear frequency-based criterion for vocabulary selection. Unlike *Anda*, *Imagina* includes over 65% of the words in band 1, almost 43% of those in band 2, and 31% of those in band 3. The book, thus, offers a balance between decent coverage and the inclusion of a high number of words among the first 3,000. Again, other textbooks, such as *En Comunidad* or *Punto y Aparte*, include more words from those first three bands, but this is only because they include more words in total, at all frequency bands.

In order to get a general idea of the number of words from each of the first three frequency bands that will be encountered by students using any of these textbooks in the first two years of college-level Spanish, a comparative analysis was carried out for the words that are shared between ETs and ITs and those that are specific to each type of book. Only the words that appeared in over half the glossaries in each category were selected for analysis. Thus, only words that appeared in at least five

of the eight ETs were included in the analyses, as were those that appeared in at least five of the eight ITs. This criterion ensures that the included words appear in most textbooks and, thus, will be encountered by most L2 Spanish students learning Spanish in American universities. Table 5.4 presents the results of this analysis.

The increase in words from the ETs to the ITs is not very high, as only 270 new words are added within the first three bands.⁸ After two years of language study, learners have been exposed to 59% of the words in band 1 and less than 33% of the words in band 2. This is not consistent with Schmitt and Schmitt's (2014) suggestion that the 3,000 most frequent words should be prioritized in language programs, as none of the first three frequency bands was completely covered in these glossaries targeting the first two years of instruction. However, the total number of words (1,619) presented in the ITs and ETs that we analyzed does approach the 2,000 words that can be learned by L2 students in two years, according to Gairns and Redman's (1986) suggestion of teaching 8–12 words per class.

Concreteness and Length in Textbooks

The unique words in ETs and ITs displayed in Table 5.4 were further analyzed to explore whether length (measured in number of letters) and concreteness (ranked from 1 to 7 by native speakers) vary by proficiency level. Table 5.5 shows the *t*-test analyses indicating that words in ITs were significantly longer and less concrete than words in ETs, thus adding to the learning burden of those new words. Length increases not only across proficiency levels but also by bands, as words in band 1 are the shortest, and low-frequency words are the longest. The trend is exactly the opposite when it comes to concreteness, with words in higher frequency bands being more concrete than those in lower bands.

Table 5.4. Unique and Shared Words per Frequency Band across all Elementary Texts and Intermediate Texts

	Unique ETs	Shared	Unique ITs	Total	Percentage of Band(s)
Band 1	182	302	106	590	59
Band 2	123	101	109	333	33.3
Band 3	89	52	55	196	19.6
Total bands 1–3	394	455	270	1,119	37.3
Low frequency	270	124	106	500	2.9*
Total	664	579	376	1,619	

* This percentage was calculated from the remaining 17,000 words in the Davies corpus.

⁸This claim is based on the analysis of all the words that appeared in at least five of the textbooks. If a word appeared in one, two, three, or four of the textbooks, but not in five of them, it was not included in the analysis.

Table 5.5. Lexicometric Characteristics of Words per Frequency Band in Elementary and Intermediate Texts

	Length Unique ETs	Length Unique ITs	t-test length	Concreteness Unique ETs	Concre -teness Unique ITs	t-test concre -teness
Band 1	5.92 (1.82)	6.71 (1.84)	3.48*	4.9 (0.99)	3.73 (0.87)	9.03**
Band 2	6.63 (2.04)	7.28 (2.26)	2.31*	5.26 (0.91)	4.23 (0.76)	8.42**
Band 3	6.57 (1.69)	7.36 (2.07)	2.51*	5.2 (0.88)	4.41 (.87)	4.5**
Low frequency	7.16 (2.22)	7.87 (2.53)	2.65*	5.67 (0.79)	4.54 (.96)	8.08**

Note: ** = $p < .001$, * = $p < .05$

Discussion

Following Schmitt and Schmitt (2014), the 3,000 most frequent words should be a main vocabulary-acquisition goal for students enrolled in a language course. In American universities, this would mean that 3,000 words should be learned during the first two years of language instruction. However, Gairns and Redman (1986) calculated that approximately 10 words could be taught per class, which amounts to 1,000 in a year and 2,000 in two years. Of course, that all these words can be taught does not mean that students will learn all of them. Thus, learning 3,000 words in two years might not be feasible in most instructional settings. In all cases, the final selection of words to include in a language course, be it 1,000, 2,000, or 3,000, should be drawn to the extent possible from the 3,000 most frequent words.

According to the results of this study, textbook authors do not seem to prioritize those 3,000 words when selecting the vocabulary to be included in their glossaries. Less than half of the 3,000 most frequent words of Spanish are generally included in ETs and ITs, and words that are among those 3,000 represent approximately 50%–64% of words in the glossaries. The other 35% of words in the textbooks are low-frequency words that may not be as useful. These results echo those obtained in previous analyses of word frequency in L2 textbooks (Davies & Face, 2006; Godev, 2009; Lipinski, 2010), which also concluded that textbook authors do not seem to base their vocabulary selection on a clear frequency criterion.

The only textbooks that included a higher number of high-frequency words were those that contained more than 3,000 words in their glossaries. This implied that they also presented more low-frequency words than the shorter glossaries and that they included more vocabulary items than those that can be expected to be learned in two years of language instruction.

Additionally, the number of words among the 3,000 most frequent increased by only 270 words from ETs to ITs, which is consistent with Alsaif and Milton's (2012) observation that L2 English textbooks at higher levels of proficiency do not add much vocabulary to that already presented in lower proficiency levels. Interestingly, the number of low-frequency words even decreased in ITs and many infrequent words were shared between ETs and ITs. This might be due to the fact that infrequent words deal with classroom topics (e.g., pizarra [blackboard] and rotulador [marker]) or contents related to grammar and metalinguistic terms (e.g., sustantivo [noun], adjetivo [adjective], oración [clause], subjuntivo [subjunctive], and gerundio [gerund]), which are used both in the first and second year of language instruction.

Overall, the situation depicted by these results might have negative implications for students, as the first 3,000 most frequent words in Spanish cover up to 94% of the words in an oral context and almost 90% of them in written texts (Davies, 2005). Limited or no exposure to these frequent words might decrease students' chances of understanding authentic written and oral texts, and learn new words from them, undermining instructors' efforts to help them become independent learners. This is even more problematic in a context where students are expected to be ready to enroll in content courses in their L2 after two years of language instruction, which would require enough vocabulary knowledge to be able to read literature texts.

Given these circumstances, it is advisable that textbook authors and publishers reevaluate the criteria they use to select vocabulary in ETs and ITs. However, instructors and LPDs cannot wait until textbooks change; they need immediate solutions. These solutions can take one or both of the following forms: (1) selecting frequent words over infrequent ones when textbooks offer long lists of words for a specific topic (e.g., targeting only the most frequent words from long lists of food items in a chapter about the supermarket) and (2) complementing classroom vocabulary exercises with homework that specifically focuses on the 3,000 most frequent words.

These solutions can be implemented easily with the aid of a frequency dictionary for the target language. For Spanish, Davies' (2006) frequency dictionary offers a helpful vocabulary list of the 5,000 most frequent words in Spanish,⁹ and similar frequency lists are available for a variety of languages (Appendix I). Additionally, the book includes lists of words, organized by frequency, for some of the most studied semantic clusters introduced in ETs and ITs, such as the vocabulary of clothing or food. Thus, an LDP could establish a clear list of vocabulary for the chapter that introduces clothing by using Davies' (2006) list of clothes to select only those items of clothing whose names are among the 3,000 most frequent words. By tailoring word lists in this way, more in-class time can be devoted to

⁹ The second edition of the *Frequency Dictionary of Spanish* (Davies & Davies, 2018) has been recently published.

those words that will be most useful for students. Without such tailoring, classroom time may be spent inefficiently in the superficial presentation of a long list of words, many of which are not used often in real-life contexts.

The vocabulary practiced in the classroom can also be supplemented with online flashcards that target words from the frequency dictionary that are not contained in the textbook. LPDs could use programs such as Quizlet, Memrise, Cerego, or Anki to develop sets of online flashcards and incorporate the study of those words into course assignments. Learners should be made aware of the goals of these learning activities and receive specific information on the importance of learning the most frequent words of their L2.

With respect to the analyses of concreteness and length, it is clear that, overall, the learning burden of words increases in ITs (see Table 5.5). Indeed, words are significantly longer and more abstract in ITs than in ETs. Interestingly, length increases incrementally with each frequency band. This offers an additional justification for focusing on the most frequent words, which are also the shortest.

Cantos and Sanchez (2011) show that the distribution of word lengths in English and Spanish is quite similar, except for 10-letter words, which appear significantly more often in Spanish than in English texts. Thus, words that are 10 letters long or more are expected to be more challenging for English speakers who learn Spanish. One exception might be the words that are 10 letters or longer but include a base that is easily identifiable by learners. For example, *pescadería* (*fish store*) has 10 letters, but students might be able to recognize *pescado* (*fish*) in it. As such, this word would not present as much of a burden as a 10-letter words such as *patrocinar*, which does not include high-frequency morphological bases. This would also be the case for long words that are cognates between Spanish and English, such as *discriminación*. Cognates would presumably not require much acquisition effort, even if this advantage is more evident for L2 learners than for L3 learners (Szubko-Sitarek, 2011). Table 5.6 presents all of the low-frequency words specific to ETs or ITs that are longer than 10 letters, specifying whether each word contains a recognizable base and whether it could be considered a cognate (i.e., either the whole word or the base resembles the English equivalent in form and meaning). The bolded words are the ones that might be more difficult to learn.

Most of the 61 low-frequency words that are 10 letters long or longer should be easy to learn, due either to their cognateness or to their recognizable base. Only eight of them (bolded in Table 5.6) would present some degree of difficulty because they are not built on frequent and transparent bases or because the semantic interpretation of their morphemic base might be confusing. For example, depending on the context, *dependiente* [*salesperson, dependent*] might not have a transparent meaning, if interpreted as a derivation of *depend* [*to depend*].

As was demonstrated above, students' ability to recognize morphemes might solve some of the issues that arise from encountering long words. Thus, it is

Table 5.6. Words 10 or more letters long in Elementary and Intermediate Texts

ETs	Cognate	Base	ITs	Cognate	Base
<i>estacionamiento</i>		X	<i>discriminación</i>	X	
<i>puertorriqueño</i>		X	<i>extraterrestre</i>	X	
<i>contabilidad</i>	X	X	<i>entendimiento</i>		X
<i>nacionalidad</i>	X	X	<i>deforestación</i>	X	
<i>refrigerador</i>	X		<i>supermercado</i>		X
<i>antibiótico</i>	X		<i>controversia</i>	X	
<i>apartamento</i>	X		<i>autorretrato</i>		X
dependiente	+/-	+/-	invernadero		
<i>despertador</i>		X	<i>desigualdad</i>		X
<i>electrónica</i>	X		<i>ascendencia</i>	X	
<i>estudiantil</i>		X	<i>emocionante</i>	X	X
<i>hamburguesa</i>	X		<i>entretenido</i>	X	X
<i>hermanastro</i>		X	<i>equilibrado</i>		X
<i>impermeable</i>	X		<i>impresionar</i>	X	X
<i>informática</i>	X		<i>medicamento</i>	X	X
mantequilla			<i>inundación</i>		X
<i>radiografía</i>	X		<i>campamento</i>		X
<i>reproductor</i>	X	X	<i>documental</i>	X	X
<i>restaurante</i>	X		<i>intermedio</i>	X	X
<i>anaranjado</i>		X	<i>analfabeto</i>	X	
<i>antipático</i>	X		<i>apasionado</i>	X	X
baloncesto			<i>autoestima</i>	X	
<i>canadiense</i>		X	<i>disponible</i>		X
<i>carpintero</i>	X		<i>entretener</i>	X	
<i>cumpleaños</i>		X	<i>inesperado</i>		X
<i>dominicano</i>		X	patrocinar		
<i>escritorio</i>		X			
estornudar					
<i>improbable</i>		X			
<i>medianoche</i>		X			
<i>microondas</i>		X			
<i>psicología</i>	X				
servilleta					
<i>sociología</i>	X				
sustantivo					

important to reflect on derivational affixation in the classroom (Sánchez-Gutiérrez, Marcos Miguel, & Robles García, in press). However, textbooks lack activities that facilitate the learning of derivation (Robles García & Sánchez-Gutiérrez, 2016; Sánchez-Gutiérrez, 2014; Neary-Sundquist, 2015). One recommendation is to identify productive affixes in the L2 and tailor activities toward them. For example, in Spanish, when teaching the vocabulary of the professions, the instructor could discuss how most words that refer to athletes or to musicians end in *-ista*. This initial presentation could trigger review activities in which the instructor starts class by asking students to recall, in one minute, as many words ending in *-ista* as they can. In such an activity, the suffix would serve as a memory retrieval cue.

For the concreteness factor, even though words in ITs are less concrete¹⁰ than those in ETs, the average concreteness rates are still high. Given that this variable is rated from 1 (extremely abstract) to 7 (extremely concrete), an average of 4–5 means that few words are extremely abstract. As such, these words do not present an added difficulty. The relative ease of learning these words can be clearly illustrated with words from Table 5.6 such as *estacionamiento*, *supermercado*, or *autorretrato*, which are low-frequency and long, but do not present the additional burden of being abstract. However, some words, even among the 3,000 most frequent, do present low rates of concreteness. Overall, verbs are rated as more abstract than nouns; thus, it would be helpful to apply a pedagogical treatment that makes verbs more imageable for the students. To this end, instructors could train students in using the keyword technique for navigating abstract vocabulary.

Conclusion

This study aimed at analyzing the frequency, length, and concreteness of the vocabulary included in Elementary Textbooks (ETs) and Intermediate Textbooks (ITs) in college-level Spanish instruction in the United States. The findings can be summarized in three points: (1) frequency is not the main factor in vocabulary selection for L2 Spanish textbooks, (2) vocabulary additions in ITs are not driven by the need to cover a certain number of frequent words, and (3) words that are exclusively presented in ITs are longer and less concrete than those that are specific to ETs. Therefore, while we clearly advocate for the more systematic inclusion of the 3,000 most frequent words in textbooks, we are aware that this change might take some time. In the meanwhile, we invite teachers and LPDs to supplement their materials so that students get the best exposure possible to those words. Several apps and programs can easily be used for this purpose, and the information about word frequency is already available for most commonly taught languages (see Appendix I). Additionally, some techniques, which previous

¹⁰ It needs to be reminded here that no function words were included in these analyses, so the difference cannot be simply due to the presence of more function words on one level over the other.

research has proven to be effective, are proposed here to address the difficulty of learning long and abstract words. Concretely, we suggest that a more systematic approach to the study of L2 words' morphological structure could be beneficial, given that it allows to interpret longer words based on morphemes that students may already recognize. We also propose that the keyword technique could be often used in class in order to address the specific challenges that abstract words present for the students. By training students in the recognition of morphological patterns and in the keyword technique at elementary levels of proficiency, we believe that instructors can contribute to lowering the burden of learning long and abstract words while also helping students to develop useful strategies that they can use on their own in more advanced levels.

References

- Alcaraz Mármol, G. (2009). Vocabulary in EFL textbooks: Frequency levels. In P. Cantos Gómez & A. Sánchez Pérez (Eds.) *A survey on corpus based studies*. Available online at <http://www.um.es/lacell/aelinco/contenido/index.html>
- Allen, H. W. (2008). Textbooks materials and foreign language teaching: Perspectives from the classroom. *NECTFL Review*, 62, 5–28.
- Alsaif, A. (2011). *Investigating Vocabulary Input and Explaining Vocabulary Uptake among EFL Learners in Saudi Arabia*. PhD dissertation, Swansea University, Swansea, UK.
- Alsaif, A., & Milton, J. (2012). Vocabulary input from school textbooks as a potential contributor to the small vocabulary uptake gained by English as a foreign language learners in Saudi Arabia. *The Language Learning Journal*, 40(1), 21–33.
- Anthony, L. (2014). AntWordProfiler (Version 1.4.1) [Computer Software]. Tokyo, Japan: Waseda University. Available online at <http://www.laurenceanthony.net/>
- Barcroft, J. (2012). *Input-Based Incremental Vocabulary Instruction*. Alexandria, VA: TESOL International Association.
- Cantos, P., & Sánchez, A. (2011). El inglés y el español desde una perspectiva cuantitativa y distributiva: equivalencias y contrastes. *Estudios Ingleses de la Universidad Complutense*, 19, 15–44.
- Criado, R., & Sánchez, A. (2009). Vocabulary in EFL textbooks: A contrastive analysis against three corpus-based word ranges. In P. Cantos Gómez & A. Sánchez Pérez (Eds.), *A survey on corpus based studies*. Available online at <http://www.um.es/lacell/aelinco/contenido/index.html>
- Davies, M. (2002) *Corpus del Español: 100 million words, 1200s–1900s*. Available online at <http://www.corpusdelespanol.org>
- Davies, M. (2005). Vocabulary range and text coverage: Insights from the forthcoming Routledge frequency dictionary of Spanish. In D. Eddington (Ed.). *Selected Proceedings of the 7th Hispanic Linguistics Symposium* (pp. 106–115). Somerville, MA: Cascadilla.
- Davies, M. (2006). *A frequency dictionary of Spanish. Core vocabulary for learners*. New York, NY: Routledge.
- Davies, M., & Davies, H. (2018). *A frequency dictionary of Spanish. Core vocabulary for learners* (2nd ed.). New York, NY: Routledge.
- Davies, M., & Face, T. L. (2006). Vocabulary coverage in Spanish textbooks: How representative is it? In N. Sagarra & A. J. Toribio (Eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium* (pp. 132–143). Somerville, MA: Cascadilla.

- De Groot, A., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign language vocabulary learning and forgetting. *Language Learning*, 50(1), 1–56.
- Donzelli, G. (2007). Foreign language learners: Words they hear and words they learn: A case study. *Estudios de Lingüística Inglesa Aplicada*, 7, 103–125.
- Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, 45(4), 1246–58.
- Ecke, P. (1999). Resumen práctico de mnemotécnicas para la enseñanza de lenguas extranjeras *Estudios de Lingüística Aplicada*, 29, 55–70.
- Ecke, P. (2004). Die Schlüsselwort-Mnemonik für den fremdsprachigen Wortschatzerwerb: Zum Stand der Forschung. *Fremdsprachen Lehren und Lernen*, 33, 213–230.
- Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, 43(4), 559–617.
- Gairns, R., & Redman, S. (1986). *Working with words. A guide to teaching and learning vocabulary*. Cambridge: Cambridge University Press.
- Godev, C. B. (2009). Word-frequency and vocabulary acquisition: An analysis of elementary Spanish college textbooks in the USA. *Revista de Lingüística Teórica y Aplicada*, 47(2), 51–68.
- Horst, M. (2013). Mainstreaming second language vocabulary acquisition. *The Canadian Journal of Applied Linguistics*, 16(1), 171–188.
- Hulstijn, J. H. (1997). Mnemonic methods in foreign language vocabulary learning: Theoretical considerations and pedagogical implications. In J. Coady & T. Huckin (Eds.), *Vocabulary acquisition: A rationale for pedagogy* (pp. 203–224). Cambridge, Cambridge University Press.
- Institute of International Education. (2015). *Open Doors Report*. IIE Books. Available online at <https://www.iie.org/en/Research-and-Insights/Open-Doors>
- Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, 319(5865), 966–968.
- Lang, M. F. (2009). *Formación de palabras en español. Morfología derivativa productiva en el léxico moderno*. Madrid: Cátedra.
- Laufer, B. (2012). Word difficulty. In C. A. Chapelle (Ed.), *Encyclopedia of Applied Linguistics*, Oxford: Wiley-Blackwell.
- Laufer, B. (1990). Why are some words more difficult than others? Some intralexical factors that affect the learning of words. *IRAL-International Review of Applied Linguistics in Language Teaching*, 28(4), 293–308.
- Lipinski, S. (2010). A frequency analysis of vocabulary in three first-year textbooks of German. *Die Unterrichtspraxis/Teaching German*, 43(2), 167–174.
- López Jiménez, M. D. (2014). A critical analysis of the vocabulary in L2 Spanish textbooks. *Porta Linguarum*, 21, 163–181.
- Marcos Miguel, N. (2015). Textbook consumption in the classroom: Analyzing a classroom corpus. *Procedia Social and Behavioral Sciences*, 198, 309–319.
- Masrai, A., & Milton, J. (2015). Investigating the relationship between the morphological processing of regular and irregular words and L2 vocabulary acquisition. *International Journal of Applied Linguistics and English Literature*, 4(4), 192–199.
- McGrath, I. (2013). *Teaching materials and the roles of EFL/ESL Teachers*. London: Bloomsbury.
- McDonough, J., Shaw, C., & Masuhara, H. (2013). *Materials and methods in ELT. A teacher's guide*. Oxford: Wiley-Blackwell.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.
- Morin, R. (2003). Derivational morphological analysis as a strategy for vocabulary acquisition in Spanish. *The Modern Language Journal*, 87(2), 200–221.

- Morin, R. (2006). Building depth of Spanish L2 vocabulary by building and using word families. *Hispania*, 89(1), 170–182.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59–82.
- Heatley, A., & Nation, I. S. P. (1994). RANGE [Computer Software]. University of Wellington, Victoria. Available online at <https://www.victoria.ac.nz/lals/about/staff/paul-nation>
- Neary-Sundquist, C. (2015). Aspects of vocabulary knowledge in German textbooks. *Foreign Language Annals*, 48(1), 68–81.
- Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research*, 20(1), 113–138.
- Robles García, P., & Sánchez-Gutiérrez, C. (2016). La morfología derivativa en los manuales de español elemental estadounidenses: un estudio exploratorio. *Revista Electrónica de Lingüística Aplicada*, 15(1), 70–86.
- Sánchez-Gutiérrez, C. (2014). Morfología derivativa y manuales de E/LE: un análisis crítico. *Anexos de la Revista Española de Lexicografía*, 22, 163–178.
- Sánchez-Gutiérrez, C., Marcos Miguel, N., & Robles García, P. (in press). What derivational suffixes should we teach in Spanish as a second language courses? *Issues in Hispanic and Lusophone Linguistics*, 16.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26–43.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484–503.
- Szubko-Sitarek, W. (2011). Cognate facilitation effects in trilingual word recognition. *Studies in Second Language Learning and Teaching*, 1(2), 189–208.
- Van Zeeland, H., & Schmitt, N. (2012). Lexical coverage and L2 listening comprehension: How much does vocabulary knowledge contribute to understanding spoken language? *Applied Linguistics*, 34(4) 457–479.
- Webb, S., & Rodgers, M. P. (2009a). Vocabulary demands of television programs. *Language Learning*, 59(2), 335–366.
- Webb, S., & Rodgers, M. P. (2009b). The lexical coverage of movies. *Applied Linguistics*, 30(3), 407–427.
- Willis, M., & Ohashi, Y. (2012). A model of L2 vocabulary learning and retention. *The Language Learning Journal*, 40(1), 125–137.

Appendix I. Resources to Find Word Frequencies

Routledge has frequency dictionaries for Spanish, Persian, Turkish, Korean, Japanese, Dutch, Russian, Arabic, Czech, French, German, Mandarin Chinese, and Portuguese.

Other corpora and frequency lists can be found online:

ARABIC

<http://arabicorpus.byu.edu/>

ENGLISH

<https://www.victoria.ac.nz/lals/resources>

<https://www.lextutor.ca/>

<https://www.wordandphrase.info/>

GERMAN

<http://wortschatz.uni-leipzig.de/en>

PORTUGUESE

<http://www.corpusdoportugues.org/>

SPANISH

<http://www.bcbl.eu/databases/espal/index.php>

<http://www.corpusdelespanol.org/>

<https://www.wordandphrase.info/span/>