

# WORD-FREQUENCY AND VOCABULARY ACQUISITION: AN ANALYSIS OF ELEMENTARY SPANISH COLLEGE TEXTBOOKS IN THE USA

FRECUENCIA LÉXICA Y ADQUISICIÓN DE VOCABULARIO:  
UN ANÁLISIS DE LOS LIBROS DE TEXTO ESTADOUNIDENSES  
DE PRIMER AÑO UNIVERSITARIO DE ESPAÑOL

---

CONCEPCIÓN B. GODEV  
University of North Carolina at Charlotte. USA.  
cgodev@uncc.edu

## ABSTRACT

At universities in the United States, first-year Spanish instructors often wonder why their students perform below expectations in vocabulary tests. This paper addresses this question by presenting an analysis of five leading first-year college Spanish textbooks published in the United States. This analysis consists of two parts: 1) mapping the textbooks' vocabulary lists as well as the vocabulary that appears in their reading input against the vocabulary-frequency information provided in *A frequency dictionary of Spanish: Core vocabulary for learners* written by Mark Davies and published by Routledge in 2006, and 2) using estimations available in the current literature to infer logistical aspects that need to be considered in first-year college Spanish such as amount of contact-time with the material to be learned, depth of vocabulary knowledge, and the relationship between the receptive and productive dimensions of vocabulary acquisition. The results show that some pedagogical fine-tuning may be needed to align the amount of vocabulary presented in textbooks as well as their vocabulary selections with research findings on vocabulary acquisition and Spanish vocabulary frequency.

*Keywords:* Frequency, vocabulary, textbooks.

## RESUMEN

En el ámbito de las universidades estadounidenses, a menudo los profesores de primer año de español constatan que los resultados de los exámenes de vocabulario están por debajo de sus expectativas. Este trabajo aborda esta cuestión por medio del análisis de cinco libros de texto publicados en los Estados Unidos. Este análisis consiste en dos partes: 1) una comparación tanto de las listas de vocabulario como del vocabulario de las lecturas con la información sobre vocabulario y frecuencia que aparece en *A frequency dictionary of Spanish: Core vocabulary for learners*, escrito por Mark Davies y publicado por Routledge

en 2006, y 2) una exposición de los aspectos logísticos que deben considerarse, tales como cantidad de tiempo en contacto con la materia y profundidad del conocimiento léxico, y de cómo se perfilan estos aspectos a través de las dimensiones receptoras y productivas del proceso de adquisición de vocabulario. Los resultados muestran que los libros de texto pueden perfeccionarse en lo que se refiere a la cantidad de vocabulario y la selección del mismo para conseguir un mejor alineamiento con los conocimientos que se tienen en este momento sobre frecuencia del léxico español.

*Palabras clave:* Frecuencia, vocabulario, libros de texto.

*Recibido:* 30-06-2009. *Aceptado:* 28-10-2009.

## 1. INTRODUCTION

Rehearsal, repetition, and frequency are all facets of time, which ultimately is the ever-present factor in language processing and language acquisition. The so-called vocabulary frequency is that aspect of vocabulary that has to do with the number of times, and therefore the duration of contact with a word, which facilitates its acquisition by enhancing the chances to draw learners' attention to it. Learners' amount of contact with words increases the chances to deepen their knowledge of a word, not only in terms of meaning but also in spelling, pronunciation, syntax and morphology<sup>1</sup>.

While there is a large body of research that provides insight regarding vocabulary frequency in English as a second language (Nation, 1990, 2001), this line of research has been very limited in the area of Spanish as a second language. However, the publication of Davies's (2006) dictionary as well as his online *Corpus del español* (Davies, 2002) are works that now make it possible for researchers to advance their knowledge on aspects of vocabulary acquisition in Spanish as a second language as they relate to vocabulary frequency. One basic question that researchers and instructors have not been able to answer until recently is how consistent is the vocabulary presented in pedagogical materials with the most highly-frequent vocabulary. The assumption underlying this question is that language teaching and learning aim at providing learners with the necessary tools to be able to communicate in real-life situations. Therefore, within this framework it seems relevant to expect that textbooks' vocabulary draw on the vocabulary that language users handle most frequently. Davies and Face (2006) carried out an analysis of the vocabulary presented in three first-year college Spanish textbooks and three second-year college Spanish textbooks. Their analysis focused on productive vocabulary, that is, vocabulary that students were expected to learn

<sup>1</sup> See Urzúa, Sáez y Echeverría (2006) for an insightful study on lexical availability in L1-Spanish.

and use. They assumed that the vocabulary lists that appeared in the chapters were productive vocabulary. They analyzed the frequency level of this vocabulary by using Davies's (2006) dictionary as a reference. The results of their study indicated that 10-50% of the words in the textbooks were among the most frequently used words.

The present study builds upon Davies and Face (2006) in that it seeks to partially replicate their analysis and answer the general question of how accurately vocabulary frequency is represented in textbooks<sup>2</sup>. The answer to this question is presented here in the light of findings on second-language acquisition regarding pace of L2 vocabulary learning in the foreign language classroom, which is one aspect that Davies and Face (2006: 142) themselves acknowledge that they did not address in their study. This paper discusses vocabulary frequency in light of findings regarding depth of vocabulary learning, the cyclical nature of vocabulary learning, and vocabulary size and reading comprehension. There are also two additional differences between the present study and the study by Davies and Face (2006); one is that while they analyzed first-year and second-year college Spanish textbooks, the present study focuses only on first-year textbooks; the other difference is that not only vocabulary lists but also vocabulary from reading input was examined in the present study.

I will present an analysis of five leading first-year college Spanish textbooks published in the United States. This analysis will consist of two parts: 1) mapping their vocabulary lists as well as the vocabulary that appears in their reading input against the vocabulary-frequency information provided in Davies's (2006) dictionary and 2) estimating logistical aspects that need to be considered in first-year Spanish, such as amount of contact-time with the material to be learned, depth of vocabulary knowledge, and the relationship of both the receptive and productive dimensions to language acquisition. The ultimate goals of this analysis are to raise instructors' awareness of the relationship between vocabulary frequencies used by native speakers and the vocabulary selections included in textbooks, and to stimulate reflection on vocabulary acquisition objectives in light of the information available regarding pace of acquisition and vocabulary frequency.

### 1.1. Assumptions

The meaning of the terms "elementary/novice Spanish" and "intermediate Spanish" in the context of college or university course offerings is likely not to

<sup>2</sup> A comparative statistical analysis between the data in Davies and Face (2006) and the data in the present study was not practicable because the frequency rankings are different in both studies. Davies and Face established five frequency rankings, whereas the present study presents three.

have a correspondence with the meaning of the same terms as used in the ACTFL<sup>3</sup> proficiency guidelines. A typical semester provides students with some forty-two hours of in-class contact with the language, provided the class meets three times a week for fourteen weeks. While students are expected to maintain contact with the subject matter, it is unlikely that they maintain an average of two hours of contact on their own for every class period, as it is explicitly recommended by some universities. According to a study that surveyed students at the University of California at Berkeley, a more realistic estimate may be that students spend an average of one hour of outside-class contact with the material for every class period (Brint and Cantwell, 2008: 2). Therefore students may have a total estimated in-class and outside-class language contact of some eighty-to-ninety hours in one semester. This kind of exposure to the language will not be sufficient for many students to move out of the novice level, as defined by ACTFL, and into the intermediate level (Hadley, 2001: 25-26).

Vocabulary learning encompasses an array of processes that interact and contribute to deepening gradually the semantic knowledge of words. These processes occur in a cyclical way as learners develop their reading, writing, listening, and speaking proficiency over time (Ellis, 2008). These processes also occur in such a way that the learning of a vocabulary item is not an all-or-nothing event. Learners may acquire a vocabulary item to the point where they may be able to recognize it in reading input but not in listening input, may be able to retrieve the item from memory while writing but not while speaking, may be able to recognize the word or retrieve the phonetic or written form but not recall its meaning, may be able to recall the meaning in the short term but not in the long term, and may be able to recognize one meaning and be unaware of other meanings (Nation, 1990: 31-32). The focus of this study is vocabulary acquisition as it relates to the development of the so-called *receptive vocabulary*, that is, vocabulary whose form and meaning the learner can recognize while reading or listening but may not necessarily be able to retrieve for the purpose of writing or speaking. This study also focuses on receptive vocabulary as it appears in reading input.

Following Laufer's (1998: 265) estimates, the present study assumes that learners learn 8-9 word families per in-class contact with the language. This assumption implies that, if in-class contact yields the learning of 8-9 word families regardless of the semester, college students who are true beginners in elementary Spanish may increase their receptive vocabulary at the rate of three hundred word

<sup>3</sup> The American Council on the Teaching of Foreign Languages (ACTFL) has established language proficiency guidelines to describe language competence levels. These guidelines are similar to those included in the Common European Framework of Reference for Languages (CEFR).

families per semester, that is, students will have a receptive vocabulary of about one thousand words after three semesters.

Another assumption in this study is that vocabulary frequency, as it happens in the target language when native speakers use it, is a criterion that merits consideration when the goal is to develop receptive vocabulary in a systematic fashion. If we accept this premise, vocabulary frequency should be a criterion for programming vocabulary instruction. If true beginners will acquire a receptive vocabulary of one thousand words over three semesters, the words that need to be given priority are those that are used most frequently in the language. ESL research suggests that these one thousand words may be particularly important if, as in English, they cover 77% of running words in academic texts (Nation, 2001: 156). Davies (2005: 109) confirms and fine-tunes this estimate for Spanish by reporting that the first one thousand most frequent Spanish words will provide a text coverage of 76% of non-fiction texts, 79.6% of fiction texts, and 87.8% of oral texts.

Vocabulary size has been found to be a predictor of reading comprehension (Anderson and Freebody, 1981; Koda, 1989; Coady, Magoto, Hubbard, Graney and Mokhtari, 1993) and has been shown to correlate to writing proficiency (Linnarud, 1986; Astika, 1993). Therefore attention to a systematic approach to vocabulary development may enhance reading comprehension.

Laufer (1992), cited in Nation (2001: 146), concluded that three thousand word families is the minimum for the reading of unsimplified texts written in English. Nation (2001: 156) also indicates that first-graders know close to five thousand word families in English, high-frequency words included. Among these words, it is assumed that the first one thousand most-frequent words play an important role in reading comprehension, because they cover about 77% of the running words in academic texts; the next one thousand covers about 5% (Nation, 2001: 16). For Spanish, Davies (2005: 109) found that the next one thousand words cover 8% of tokens in a non-fiction text, 6.5% of those in a fiction text, and 5% of tokens in an oral text.

Another aspect to vocabulary proficiency and reading is what Laufer labels as “sight vocabulary,” that is, “words whose forms and common meanings are recognized automatically, irrespective of context” (1997: 22). This is another dimension of frequency that suggests that learners could potentially benefit from planned instruction that is consistent with vocabulary frequencies as they occur in the target-language oral and written texts. The link between sight vocabulary and vocabulary frequency lies in the theoretical assumption that a large component of the sight vocabulary is made up of the most frequent vocabulary (Huckin and Coady, 1999: 184).

## 2. METHOD

### 2.1. Materials

The data analyzed in this study come from the following sources<sup>4</sup>:

1. *A frequency dictionary of Spanish: Core vocabulary for learners* by Davies (2006).
2. A sample of three hundred and eighty-four words from the vocabulary at the end of the textbook, end-vocabulary hereafter, and Chapter 2 (reading) in *Puntos de partida* (8<sup>th</sup>) by Knorre, Dorwick, Pérez Gironés, Glass and Villarreal (2009). The reading is a one-hundred-and-eleven-word text, from page ninety-three, about families in the US and Canada. See summary in Table I.
3. Chapters 2 (vocabulary) and 10 (reading) from *Tú dirás* (4<sup>th</sup>) by Martínez-Lage, Gutiérrez and Rosser (2007). The reading is a sixty-nine-word poem from page four hundred and eleven. See summary in Table I.
4. Chapters 4 (vocabulary) and 7 (reading) from *Dicho y hecho* (8<sup>th</sup>) by Dawson, Potowski and Sobral (2008). The reading is a one-hundred-and-three-word text from page two hundred and twenty-four, that provides general information about Chile. See summary in Table I.
5. Chapters 5 (vocabulary) and 4 (reading) from *Arriba* (5<sup>th</sup>) by Zayas-Bazán, Bacon and Nibert (2008). The reading is a one-hundred-and-seventy-one-word text, from page one hundred and thirty-one, that provides information about the typical Hispanic family. See summary in Table I.
6. Chapters 13 (vocabulary) and 9 (reading) from *Nexos* (1<sup>st</sup>) by Spaine Long, Carreira, Madrigal Velasco and Swanson (2007). The reading is a one-hundred-and-forty-word text, from pages two hundred and eighty-eight and two hundred and eighty-nine, about Bolivian traditions. See summary in Table I.

<sup>4</sup> This study should not be interpreted as an endorsement or rejection of any of the texts analyzed, but rather only as a body of evidence for research.

**Table I.** Summary of the textbook material analyzed.

	Textbooks				
	Puntos de partida	<i>Tú dirás</i>	<i>Dicho y hecho</i>	<i>Arriba</i>	<i>Nexos</i>
end-of-textbook vocabulary	384-word sample				
chapter vocabulary list		vocabulary list in chapter 2	vocabulary list in chapter 4	vocabulary list in chapter 5	vocabulary list in chapter 13
reading	111-word reading in chapter 2	69-word reading in chapter 10	103-word reading in chapter 7	171-word reading in chapter 4	140-word reading in chapter 9

## 2.2. Sampling procedure

The end-vocabulary three-hundred-and-eighty-four-word sample comes from three pages that were randomly selected from a total of thirty-three pages that contain four thousand five hundred and twenty-four words. These pages are at the end of *Puntos de partida* and they include the vocabulary that appears in the vocabulary sections in all the chapters. The purpose of the analysis of this sample is to find out whether frequency patterns in the eighth edition were similar to those in the sixth edition, which was the edition analyzed in Davies and Face (2006).

The entire four-thousand-five-hundred-and-twenty-four-word end-vocabulary in *Puntos de partida* was analyzed to verify whether the first one thousand most-frequent words in Davis (2006) were included.

One chapter from three of the textbooks was randomly selected from the first ten chapters, which are the chapters usually covered in the first two semesters of college Spanish. The frequency of the vocabulary that appears at the end of these chapters (end-of-chapter vocabulary hereafter) was analyzed.

One chapter from the last five of *Nexos*, a chapter covered in second year, was randomly selected in order to examine how the frequency of the vocabulary at the end of this chapter compared to that of chapters covered in first year.

The five reading selections that were analyzed come from one chapter from each of the five textbooks listed above. Each chapter was randomly selected out of the first ten chapters in each of the textbooks.

## 2.3. Analysis

The vocabulary from the textbooks was lemmatized according to the procedure explained in Davies (2005: 109) and Davies and Face (2006: 136).

Following the assumption mentioned earlier that students may learn three-hundred word families in a given semester and the most frequent five thousand words in Davis (2006), the end-vocabulary, the end-of-chapter vocabularies, and the reading vocabulary (i. e., the vocabulary that was in the reading selections) were grouped into three frequency ranges, namely, one through one thousand, one thousand through five thousand, and above five thousand. The frequency ranges adopted in this study reflect the assumption that the focus of college first-year Spanish learning may need to be on the first one thousand most frequently used words if students can only learn some three hundred word families in one semester, as was indicated in section 1.1. Therefore any word falling in the frequency ranking of one-thousand-and-one-to-five-thousand, which is the target for intermediate level Spanish, are classified in a separate range. Finally, any word above the intermediate level is categorized in the above-five-thousand range.

The frequency ranking of each word was determined in reference to the ranking assigned to them by Davies (2006). The lower the frequency ranking number the more frequent the word, e. g., among the five thousand most frequently used words in Spanish, Davies (2006) ranks the definite article as number one, the most frequent word, and the word *cueva* 'cave' as number five thousand, the least frequent word. The words from the textbook that did not appear in Davies (2006) were assumed to fall in a frequency range above the five thousand ranking.

The vocabulary in the reading selections was analyzed to determine whether or not students would be likely to understand it with a reasonable level of accuracy. This evaluation was based on the following criteria: 1) the item was a cognate that could be easily associated with its English counterpart, 2) the item had already been introduced in the textbook by the time students encountered the reading, and 3) the item was encountered in the textbook more than once.

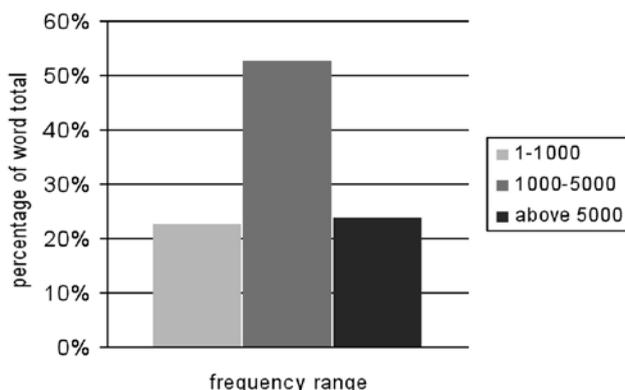
The vocabulary in the reading was also analyzed to determine the frequency range of that part of the vocabulary that was determined to be unfamiliar to the students.

### 3. RESULTS

*Puntos* end-vocabulary includes a total of four thousand five hundred and twenty-four words. The *Puntos* end-vocabulary sample of three hundred and eighty-four words analyzed here is made up of 23% of words ranked within the one-to-one-thousand frequency range, 53% of the sample falls in the one-thousand-to-five-thousand frequency range, and 24% falls in the above-five-thousand range. See Table II and Figure 1. Forty-four words among those ranked in Davies (2006) as the one thousand most frequently used words are not listed in the *Puntos* four-thousand-five-hundred-and-twenty-four-word end-vocabulary as a whole (see Appendix).

**Table II.** *Puntos* (8<sup>th</sup>) end-vocabulary three-page sample. Percentage of words in three different frequency ranges.

		Frequency range					
		1-1000		1000-5000		above 5000	
n	%	n	%	n	%	total	
89	23	203	53	92	24	384	



**Figure 1.** *Puntos* end-vocabulary sample.

The end-of-chapter vocabulary from the four chapters includes 25, 30, 27 and 27% of words ranked within the one-to-one-thousand frequency, 52, 32, 41 and 60% of words within the one-thousand-to-five-thousand frequency, and 23, 38, 32 and 13% of words within the above-five-thousand frequency. See Table III and Figure 2.

**Table III.** Percentage of end-of-chapter vocabulary in three different frequency ranges.

	Frequency range						total	
	1-1000		1000-5000		above 5000		n	%
	n	%	n	%	n	%	n	%
Tú dirás ch2	31	25	66	52	29	23	126	100
Dicho ch4	35	30	37	32	45	38	117	100
Arriba ch5	20	27	30	41	23	32	73	100
Nexos ch13	22	27	48	60	10	13	80	100

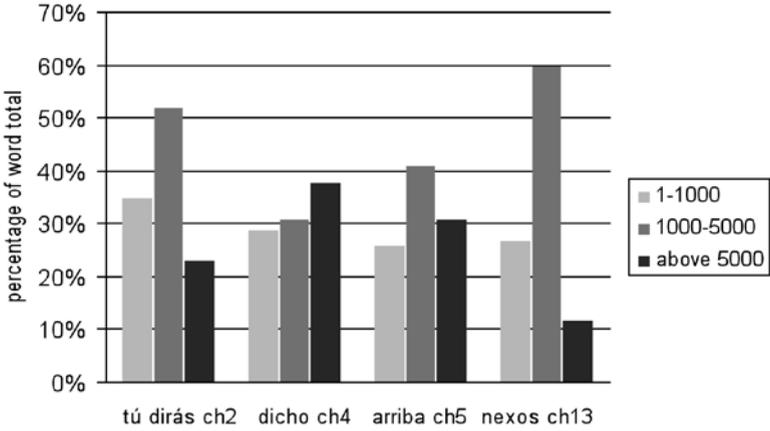


Figure 2. End-of-chapter vocabulary.

The selected readings were found to include unknown words estimated in the following percentages: 22, 17, 22, 11 and 27%. The percentage of glossed words out of the word total is 9, 17, 3, 0 and 12% respectively. The text coverage, defined as the number of words known, is estimated as 78, 83, 78, 89 and 73% respectively. See Table IV and Figure 3.

Table IV. Coverage and glosses in reading selections.

	unknown words		glossed words		coverage		word total in selection
	n	%	n	%	n	%	
Puntos ch2	25	22	10	9	86	78	111
Arriba ch4	29	17	10	17	142	83	171
Dicho ch7	23	22	3	3	80	78	103
Nexos ch9	15	11	0	0	125	89	140
Tú dirás ch10	19	27	8	12	50	73	69

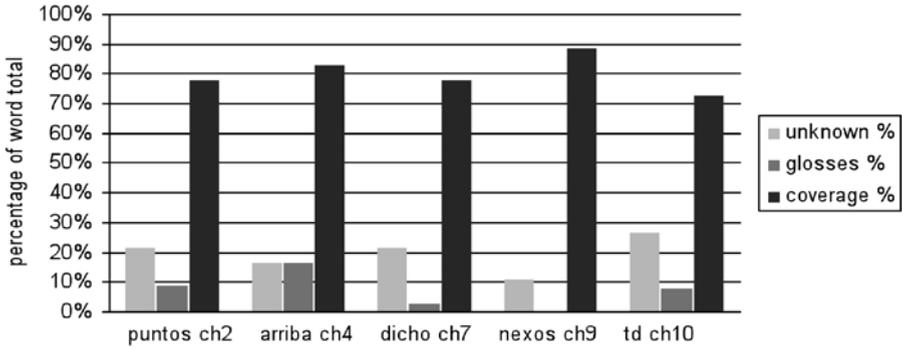


Figure 3. Reading coverage.

The frequency ranges of these unknown words are as follows: 68, 79, 56, 40 and 32% of unknown words are in the one-to-one-thousand frequency range, 16, 17, 22, 7 and 26% are in the one-thousand-to-five-thousand frequency range and 16, 4, 22, 53 and 42% are in the above-five-thousand frequency range. See Table V and Figure 4.

Table V. Unknown reading vocabulary and frequency range.

	Frequency Range						total	
	1-1000		1000-5000		above 5000		n	%
Puntos ch2	n	%	n	%	n	%	25	100
Arriba ch4	23	79	5	17	1	4	29	100
Dicho ch7	13	56	5	22	5	22	23	100
Nexos ch9	6	40	1	7	8	53	15	100
Tú dirás ch10	6	32	5	26	8	42	19	100

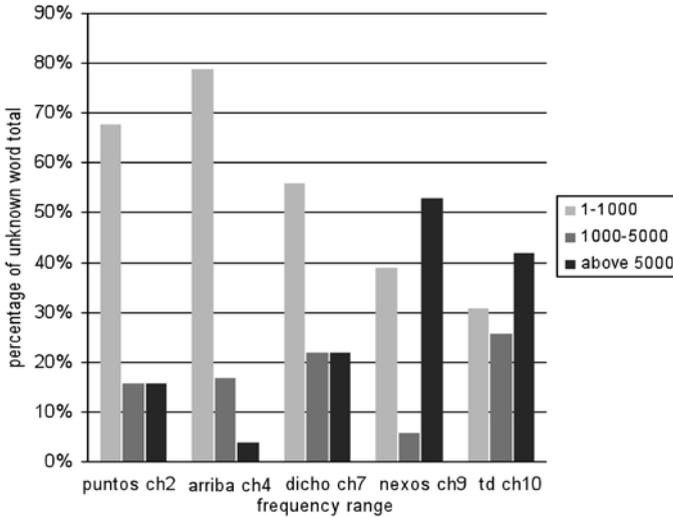


Figure 4. Unknown reading vocabulary.

#### 4. DISCUSSION

The results of the analysis of *Puntos* end-vocabulary partially replicate the results that Davies and Face (2006: 140) report. The differences between the results of the present study and theirs can be attributed to the fact that they analyzed the sixth edition and the entire end-vocabulary while in the present study the data came from a sample from the eighth edition. They also examined the data by breaking down frequency ranges in a different fashion, ten different ranges in their study vs. three in this one.

*Puntos* end-vocabulary sample in this study featured 23% of vocabulary within the one-to-one-thousand frequency, whereas Davies and Face (2006: 140) report 46%. While the numerical results are different, both results reveal that a large number of vocabulary items in this first-year Spanish textbook falls out of the one-to-one-thousand most frequent range that may ideally need to be targeted. In terms of targeting the most frequent words, the result that stands out is that 24% of the vocabulary in the sample was in the above-five-thousand frequency range.

The fact that forty-four words out of the one thousand most frequent words listed in Davies (2006) are not included in *Puntos*'s four-thousand-five-hundred-and-twenty-four-word end-vocabulary is an indication that vocabulary selections in first-year Spanish textbooks may need some fine-tuning. Within the context of a first-year textbook, it seems counterintuitive to overlook forty-four words that are within the targeted frequency range while many other words within a higher range

are included. Davies and Face (2006: 140) have already reported that many of the words within the one-to-one-thousand frequency range were altogether missing from the first-year and second-year Spanish textbooks that they had examined. Moreover, the size of the end-vocabulary does not conform with the learnability index for receptive vocabulary suggested in the literature, much less for productive vocabulary. The large number of words that falls outside the target range diverts the attention that could be better channeled towards cementing the development of the one thousand most frequently used words.

The end-of-chapter vocabularies from *Tú dirás*, *Dicho*, and *Arriba* present patterns similar to those observed in *Puntos* end-vocabulary in terms of the percentage of words in the one-to-one-thousand frequency range vs. the large number of words in the one-thousand-to-five-thousand frequency range as well as those in the above-five-thousand range. The number of words in the above-five-thousand frequency range is particularly striking in *Dicho* (ch 4) and *Arriba* (ch 5) as these are chapters that are covered in the first semester, yet the number of words in the above-five-thousand range is larger than those in the one-to-one-thousand range. In contrast, *Nexos* (ch 13), a chapter usually studied in the third semester, features the lowest number of words in the above-five-thousand range. Since *Nexos* (ch 13) is covered at a more advanced level, one would expect that, if any words above the five thousand range were to be included, there would be more of them in chapters covered in more advanced levels than in those covered in lower levels.

About 60% of the ten thousand most frequent English vocabulary comes from Latin or French (Minkova and Stockwell, 2006: 467). Since Spanish is a Romance language, we should expect that a significant number of Spanish words can readily be understood by the learner when reading. However, in spite of the vocabulary overlap between the two languages, we need to regard some of the words with caution, as the frequency of these words may differ in both languages. For instance, the word *cotidiano* is far more frequent in Spanish than its English counterpart ‘quotidian’. The word *concluir* in a reading passage may readily be associated with *to conclude* provided learners are not cognitively loaded with many other words whose meaning they do not know. Hu and Nation (2000: 419) argue that somewhere between 80 and 90% vocabulary coverage is the minimum to have adequate comprehension; they also note that at 80% coverage “skills and background knowledge cannot usually make up for lack of vocabulary knowledge”, and it is likely that words like *concluir* may not be associated with its English counterpart ‘conclude’ because of task overload.

With regards to the reading coverage (Table IV and Figure 3), in *Puntos*, the text coverage, that is, the percentage of words known to the learner, is estimated to be 78%, that is, one in every five words is not known. Something similar happens with the reading in *Dicho* (78% coverage), *Tú dirás* (73% coverage),

*Arriba* (83% coverage) and *Nexos* (89% coverage). Therefore, according to Nation (2001), these texts cannot be considered to provide meaning-focused input or fluency-focused input to promote vocabulary acquisition or fluency respectively. Nation (2001: 148) indicates that “learners would need to have 95% coverage for learning vocabulary from meaning-focused input, and 98-100% coverage for fluency development. This means that learners need to have simplified reading material of various levels in order to learn from meaning-focused input and to develop fluency in reading if they are to learn from these strands at all stages of their second language development”.

Carver (1994) points out that learners will increase their vocabulary size through reading only with material that is not too easy so that they have opportunities to encounter new words. Reading easy texts may serve the purpose of deepening vocabulary knowledge, but not the purpose of increasing vocabulary size. In this respect, the reading selections provide opportunities to encounter new vocabulary, but what they lack is an appropriate pedagogical treatment to aid the acquisition of new vocabulary. The pedagogical treatment associated with the readings in the textbooks examined here focuses on comprehension of general ideas but does not include manipulation of vocabulary.

One final aspect to this discussion is to consider the nature of the unknown vocabulary that was featured in the readings (see Table V and Figure 4). The unfamiliar vocabulary consists of 32%, in the best case, all the way to 79%, in the worst case, of words that fall within the one-to-one-thousand frequency range. In two of the readings, *Nexos* (ch 9) and *Tú dirás* (ch 10), the percentage of unfamiliar words within the above-five-thousand range was higher than the percentage of words within the one-to-one-thousand range. These results suggest that students may have more difficulty with reading comprehension than instructors may anticipate in first-year Spanish.

## 5. CONCLUSION

Provided that the vocabulary acquisition goal for receptive vocabulary should be around one thousand words, and furthermore that these should also be the most frequently used one thousand words, the elementary textbooks included in this analysis do not appear to have been designed according to this guiding principle. The most apparent evidence of this finding is that the *Puntos* end-vocabulary includes some four thousand five hundred words, 4.5 times as many words over the estimated one thousand that elementary Spanish students can acquire over two or three semesters in a three contact-hour schedule. The size of the end-of-chapter vocabularies, which include one hundred and twenty-six, one hundred and seventeen, seventy-three, and eighty items respectively, implies that students will

learn new vocabulary items at the rate of thirteen-to-twenty-one per contact hour, which is inconsistent with Laufer's (1998: 265) estimates of 8-9 per contact hour. The purpose of including so much vocabulary in first-year Spanish textbooks may be to expose learners to it without the expectation of some degree of acquisition within the first year. If this is the case, future research may need to determine whether the presence of this vocabulary is beneficial or detrimental to focusing first-year Spanish students' attention to the target vocabulary, that is, the one thousand most frequently used words.

The presence of a sizeable number of words, 24% of the four thousand five hundred and twenty-four words in the end-vocabulary, in the above-five-thousand frequency range further underlines the need for a vocabulary that is more aligned with second language acquisition principles.

At the elementary level, the approach to readings should go beyond extracting the gist of the information to include deep processing activities that enhance vocabulary acquisition, and the target vocabulary should be in the one-to-one-thousand frequency range.

The vocabulary coverage featured in the reading selections suggests that instructors need to be mindful that reading strategies and use of background knowledge may not make up for lack of vocabulary knowledge when coverage is around 80% (Hu and Nation, 2000).

Future research needs to examine the role of cognates in vocabulary acquisition. Cognates may give instructors and textbook writers a skewed sense of how familiar students of Spanish may be with Latin-based English words that are similar to Spanish, especially when instructors and writers are native speakers of Spanish. Two additional aspects of cognates is their frequency range in Spanish compared to English and whether or not the English counterparts feature similar collocations. Finally, the vocabulary identified by this author as unfamiliar to the students needs to be tested in the classroom to verify the accuracy of this author's assessment.

## REFERENCES

- Anderson, R. C. and Freebody, P. 1981. "Vocabulary knowledge". In J. Gutne (ed.) *Comprehension and teaching research reviews*. Newark, DE: IRA, pp. 77-117.
- Astika, G. G. 1993. "Analytical assessment of foreign students' writing", in *Regional language center (RELC) journal* 24, pp. 61-72.
- Brint, S. and Cantwell, A. M. 2008. "Undergraduate time use and academic outcomes: Results from UCUES 2006", *Research & Occasional Paper Series: CSHE* 14.08. [On line]. Available at <http://cshe.berkeley.edu/publications/docs/ROPS-Brint-TimeUse-9-24-08.pdf>. [Consulted: 05/ 25/2009].

- Carver, R. P. 1994. "Percentage of unknown vocabulary words in text as a function of the relative difficulty of the text: Implications for instruction", in *Journal of reading behaviour* 26, pp. 413-437.
- Coady, J., Magoto, J. Hubbard, P., Graney, J. and Mokhtari, K. 1993. "High frequency vocabulary and reading proficiency in ESL readers" In T. Huckin. M. Haynes and J. Coady (eds.) *Second language reading and vocabulary learning*. Norwood, NJ: Ablex Publishing Corporation, pp. 217-228.
- Davies, M. 2002. *Corpus del español*. [On line]. Available at <http://www.corpusdelespanol.org>. [Consulted: 06/05/2009].
- 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 7<sup>th</sup> Hispanic linguistics symposium*. Somerville, MA: Cascadilla, pp. 106-115.
- Davies, M. 2006. *A frequency dictionary of Spanish: Core vocabulary for learners*. New York, NY: Routledge.
- Davies, M. and Face T. L. 2006. "Vocabulary coverage in Spanish textbooks: How representative is it?". In N. Sagarra and A. J. Toribio (eds.) *Selected proceedings of the 9<sup>th</sup> Hispanic linguistics symposium*. Somerville, MA: Cascadilla, pp. 132-143.
- Dawson, L., Potowski, M., K. and Sobral, S. 2008. *Dicho y hecho*, 8<sup>th</sup> edition. Hoboken, NJ: John Wiley & Sons.
- Ellis, N. C. 2008. "The dynamics of second language emergence: Cycles of language use, language change, and language acquisition", in *Modern language journal* 92 (2), pp. 232-249.
- Hadley, A. O. 2001. *Teaching language in context*, 3<sup>rd</sup> edition. Boston, MA: Heinle & Heinle.
- Hu, M. and Nation, I. S. P. 2000. "Unknown vocabulary density and reading comprehension", in *Reading in a foreign language* 13 (1), pp. 403-430.
- Huckin, T. and Coady J. 1999. "Incidental vocabulary acquisition in a second language", in *Studies in second language acquisition* 21 (2), pp. 181-183.
- Knorre, M., T. Dorwick, Pérez Gironés, A. M., Glass, W. R. and Villarreal, H. 2009. *Puntos de partida*, 8<sup>th</sup> edition. New York, NY: McGraw-Hill.
- Koda, K. 1989. "The effects of transferring vocabulary knowledge on the development of L2 reading proficiency", in *Foreign language annals* 22, pp. 529-40.
- Laufer, B. 1992. "How much lexis is necessary for reading comprehension?". In H. Bejoint and P. Arnaud (eds.) *Vocabulary and applied linguistics*. Hampshire: Macmillan, pp. 126-132.
- Laufer, B. 1997. "The lexical plight in second language reading". In J. Coady and T. Huckin (eds.) *Second language vocabulary acquisition*. New York, NY: Cambridge University Press, pp. 20-34.

- Laufer, B. 1998. "The development of passive and active vocabulary in a second language: Same or different?", in *Applied linguistics* 19 (2), pp. 255-271.
- Linnarud, M. 1986. *Lexis in composition: A performance analysis of Swedish learners' written English*. Malmö: CWK Gleerup.
- Martínez-Lage, A., Gutiérrez, J. and Rosser, H. 2007. *Tú dirás*, 4<sup>th</sup> edition. Boston, MA: Heinle & Heinle.
- Minkova, D. and Stockwell, R. 2006. "English words". In B. Aarts and A. McMahon (eds.) *The handbook of English linguistics*. Oxford: Blackwell Publishing, pp. 461-482.
- Nation, I. S. P. 1990. *Teaching & learning vocabulary*. New York, NY: Newbury House.
- Nation, I. S. P. 2001. *Learning vocabulary in another language*. New York, NY: Cambridge University Press.
- Spaine Long, S., Carreira, M., Madrigal Velasco, S. and Swanson, K. 2007. *Nexos*. New York, NY: Houghton Mifflin.
- Urzúa, P., Sáez, K., Echeverría, M.S. 2006. "Disponibilidad léxica matemática: Análisis cuantitativo y cualitativo", in *Revista de lingüística teórica y aplicada* 44 (2), pp. 59-76.
- Zayas-Bazán, E., Bacon, S. and Nibert, H. 2008. *Arriba*, 5<sup>th</sup> edition. Upper Saddle River, NJ: Prentice Hall.

## APPENDIX

	Frequency rank (Davies, 2006)	Words from Davies (2006) not listed in the 4524-word end-vocabulary in <i>Puntos de partida</i> (8th)
1	193	idea <i>idea</i>
2	305	suponer <i>to suppose, assume</i>
3	333	a medida que <i>as</i>
4	363	cumplir <i>to fulfil</i>
5	366	pesar <i>sorrow</i>
6	386	pie <i>base</i>
7	387	valer <i>cost</i>
8	402	mover <i>to move, incite</i>
9	406	suceder <i>to happen</i>
10	477	cabo <i>end, bit</i>
11	493	señalar <i>to point out, signal</i>
12	530	carácter <i>personality, nature</i>
13	573	meter <i>to put into</i>
14	583	pretender <i>to attempt</i>
15	590	personal <i>personal</i>
16	595	acordar/se <i>to remember</i>
17	611	lanzar <i>to throw, launch</i>
18	616	apoyar <i>to support, lean on</i>

	Frequency rank (Davies, 2006)	Words from Davies (2006) not listed in the 4524- word end-vocabulary in <i>Puntos de partida</i> (8th)
19	641	imponer <i>to impose</i>
20	642	obligar <i>to obligate, force</i>
21	657	golpe <i>hit, strike, punch</i>
22	661	simple <i>simple, mere, simple-minded</i>
23	664	normal <i>normal. usual, regular</i>
24	667	bien <i>goods, property, benefit</i>
25	679	pleno <i>complete, full</i>
26	690	dispuesto <i>willing ready</i>
27	709	popular <i>popular</i>
28	753	salvar <i>to rescue</i>
29	777	mas <i>but, however</i>
30	785	marcha <i>march, progress</i>
31	789	pensamiento <i>thought, thinking</i>
32	847	intervenir <i>to intervene</i>
33	862	enfrentar <i>to confront, face</i>
34	887	advertir <i>to notice</i>
35	891	transformar <i>to transform</i>
36	903	diario <i>newspaper</i>
37	910	plantear <i>to propose, present</i>
38	929	medir <i>measure</i>
39	937	atravesar <i>to cross</i>
40	959	tardar <i>to delay</i>
41	964	convenir <i>to be agreeable</i>
42	969	alumno <i>student</i>
43	975	superar <i>to overcome</i>
44	994	seguramente <i>surely</i>

Forty-four words out of the one thousand most-frequently used words do not appear in *Puntos's* end-vocabulary. One word falls in the one-to-three-hundred frequency range, fifteen fall in the three-hundred-to-six-hundred frequency range, and twenty-eight fall in the six-hundred-to-one-thousand frequency range.