Title

Chapter 10 Vocabulary in the L2 Spanish classroom: What students know and what their instructors believe they know

ISBN

9783110730500

Authors

Sánchez-Gutiérrez, Claudia Helena

Álvarez, César Hoyos

Robles-García, Pablo

Publication Date

2023-09-18

DOI

10.1515/9783110730418-010

Vocabulary in the L2 Spanish classroom: What students know and what their instructors believe they know

Claudia Helena Sánchez-Gutiérrez
Associate Professor
University of California, Davis
One Shields Avenue. Davis, California, 95616
chsanchez@ucdavis.edu

Pablo Robles-García
Assistant Professor
University of Toronto, Mississauga
3359 Mississauga Rd. Mississauga, Ontario, L5L 1C6
p.roblesgarcia@utoronto.ca

César Hoyos Álvarez
PhD candidate
University of California, Davis
One Shields Avenue. Davis, California, 95616
cehoyosalvarez@ucdavis.edu

1 Introduction

Second language (L2) teachers need to decide on a daily basis which specific words to teach in their classes, and this task, though trivial at first sight, is far from being easy. Contrary to the clear grammatical learning goals that can be found in L2 course syllabi and textbooks, vocabulary goals tend to remain vague. For example, mastering the communicative task of *sharing a story from your past* would require that students learn the uses of both the imperfect and the preterit tenses. This makes it easy to establish a set of clearly defined grammatical goals that will allow students to fulfill the task. For this same task, however, which are the specific words that would allow students to complete it? The answer is not as obvious, as it will depend on the story each student wants to share. A student who wants to tell a story about their family may find the word *abuela [grandma]* extremely useful, while a student who would want to share one of their sport team's successes, may need to learn the word for *baseball bat* in Spanish. Even in cases where a task requires a specific thematic set of words, such as *buying food at the supermarket*, many questions arise. After all, is the word *zucchini* more relevant than *eggplant* when it comes to completing the task? Or is it the other way round? If we were to ignore this question, and all the similar ones that would necessarily come about, and we decided to include all the food-related words we could recall, the glossary for that particular task could amount to hundreds, if not thousands, of words, thus setting an unachievable goal for L2 learners. Due to the virtually limitless nature of any language's lexicon and the limited class time available to teach these many words, making choices about which words to teach and which to dismiss is necessary but challenging.

When making those decisions, instructors could rely on lexical frequency information, a criterion that has been acknowledged and promoted by the research community for decades now. Indeed, the 3,000 most frequent words in a language provide approximately 95% of vocabulary coverage in most oral and written texts (Davies 2005; Nation 2006; Schmitt and Schmitt 2014), which has been set as the minimal threshold for acceptable comprehension (Laufer 1989; Laufer and Ravenhorst-Kalovski 2010; Van Zeeland and Schmitt 2013), and represent the cut-off point for most L2 learner dictionaries (Schmitt and Schmitt 2014). Consequently, vocabulary researchers often invite language practitioners to prioritize this high-frequency vocabulary during the early stages of language instruction (Horst 2013; Meara 1980; Nation 2013; Stæhr 2008). In the *zucchini* vs. *eggplant* case, *eggplant* has 56 occurrences in SUBTLEX-US (Brysbaert and New 2009), a corpus of subtitles in US English, as opposed to the 49 of *zucchini*. Following a purely frequency-based decision-making process, *eggplant* should be prioritized over *zucchini*.

At this point, though, one may wonder how these research-based recommendations are followed in real classes and whether students are adequately exposed to high-frequency words in L2 courses. Namely, are students more exposed to *eggplant* than *zucchini* in their L2 classes due to the higher frequency of the former? In the absence of large-scale generalizable corpora of teacher talk in L2 classrooms, textbooks can function as proxies of actual vocabulary use in the classroom and offer insights into learners' lexical exposure, as they are "powerful indicators of prevalent pedagogical paradigms, and the content analysis of published titles can provide a snapshot of established practices and perspectives on language teaching and learning" (Cubillos 2014: 206). In terms of high-frequency vocabulary in L2 textbooks, Lipinski (2010) found that out of all the words in the glossaries of the three L2 German textbooks that she analyzed, only 32% pertained to the first 1,000 most frequent words, 17% to the second 1,000 frequency band, and 10% to the third. More than 35% of the words introduced in these textbooks did not fall within the 4,000 most frequent words. Similar findings have been reported in English (Alcaraz Mármol 2009; Criado and Sánchez 2009; Donzelli 2007; Martini 2012; Sun and Dang 2020; Yang and Coxhead 2020) and Spanish (Davies and Face 2006; López Bastidas and Sánchez-Gutiérrez 2020). Recently, Sánchez-Gutiérrez, Marcos Miguel and Olsen (2019) found that the 16 Spanish L2 textbooks that they studied included, on average, only 27% of the words from the first 1,000-word band, 17% from the second 1,000-word band, and 11% from the third.

Considering the vagueness of lexical learning goals in the syllabi and the fact that textbook glossaries do not systematically favor high-frequency words over less frequent ones, teachers are ill-equipped and left alone when it comes to determining which words to teach in their classes (McCrostie 2007). In this context, two sources of information can drive their vocabulary selection decisions: (1) lexical frequency data that they would need to retrieve from corpora or frequency lists, and (2) their own intuition. So far, research indicates that teachers' decision-making relies mostly on the latter instead of on empirical data (Creighton 2007; Earl and Katz 2006; Sánchez-Gutiérrez, Robles-García, and Pérez Serrano 2022; Vanlommel et al. 2017), and that these intuitions are primarily based on teachers' experiences and subjective feelings of knowledge (Epstein 2008). Since teachers' intuitions play such a crucial role in their pedagogical decisions, there is a logical need to assess their accuracy with regard to lexical

frequency, an aspect that is so critical in L2 vocabulary teaching/learning. Basically, do teachers know that *eggplant* is more frequent than *zucchini*?

Most research on native speakers' accuracy in determining word frequency is based on comparisons with non-native speakers (Aizawa, Mochizuki, and Meara 2001; Alderson 2007; McCrostie 2005; Ringeling 1984; Schmitt and Dunham 1999), but few studies have specifically focused on L2 teachers. To the best of our knowledge, only McCrostie (2007) has analyzed L2 English teachers' assessments of word frequency to date. In this two-fold study, the author (1) compared the word frequency judgments of 21 EFL teachers and 11 native English speakers who were not involved in language teaching, and (2) examined both groups' abilities to estimate the frequency ranks of high-frequency (i.e., words among the 2,000 most frequent) and mid frequency words (i.e., words in the 4,000 to 10,000 frequency bands). Results showed that teaching experience did not play a significant role in word frequency judgements and that participants were all much more accurate when ranking words at the extremes of the frequency bands (i.e., the most frequent and infrequent ones), while displaying notable difficulties when classifying mid frequency words.

These results suggest that, even though teachers have relatively accurate intuitions when it comes to very high frequency words, such intuitions rapidly become insufficient when moving beyond the first 2,000 most frequent words. Therefore, given the importance of teaching mostly frequent words (but not only the first 2,000), the limited guidance in syllabi, and the excessively large amount of low frequency words presented in L2 textbooks, it would be advisable for teachers to complement their intuitions with objective data from frequency corpora when selecting vocabulary for teaching purposes (McCrostie 2007). However, as was mentioned earlier, teachers do rely more on their intuition than on corpus-data, a fact that is further evidenced in Dang and Webb (2020). In their survey of 16 Vietnamese L2 English teachers, corpus-based wordlists, research-based vocabulary tests, and lexical profilers (all grounded in corpus-based computations of frequency) were considered the least useful tools when deciding which words to teach. Conversely, language instructors listed textbooks and their own linguistic experience as their main sources of information for vocabulary selection. These data confirm the complicated situation currently in place: language teachers do not actively use corpus-based lexical frequency data due to lack of familiarity with the tools or overall distrust for such data. Furthermore, their own intuitions are not to be trusted either when it comes to word frequencies past the 2,000 most frequent, and the textbooks they use do not offer a systematic coverage of high-frequency words.

In an interesting turn of events, though, recent literature has advocated for vocabulary wordlists creation practices that transform teachers' role from passive "consumers" of corpus-based frequency lists —which they are not anyway— to active contributors to wordlists that better address the needs of students in real classrooms (He and Godfroid 2019; Stein 2017). While frequency lists, if actually used, have the potential of helping teachers identify words that provide greater lexical coverage in most texts, recent studies have demonstrated that this information may not be directly transferable into the classroom unless *usefulness* is taken into consideration (Dang, Webb, and Coxhead 2020; Garnier and Schmitt 2015; Laufer and Nation 2012). The introduction of this additional criterion may contribute to closing the distance

between researchers, who focus on objective corpus-based measures to determine whether a word needs to be taught or not, and teachers, who base those decisions on their daily experience with actual students in the classroom. In addition to this call for a greater involvement of teachers in the creation of wordlists, through their providing usefulness ratings, other authors have also advocated for a more bottom-up approach in establishing the lists of words that should be taught at different proficiency levels. Concretely, Brysbaert, Keuleers, and Mandera (2020) have recently published vocabulary lists based on the words that real L2 English students actually know, thus establishing current students' word knowledge as a criterion to decide which words should be taught at each proficiency level to future students.

In a similar line of inquiry, which unites the calls for both teachers' and instructors' experiences to be considered when creating wordlists, Dang, Webb, and Coxhead (2020) examined the usefulness of different wordlists as evidenced in teachers' perceptions and students' actual vocabulary knowledge. A total of 973 non-overlapping headwords were selected from two different corpora: 545 from the Nation's (2012) BNC/COCA2000 (British National Corpus/Corpus of Contemporary American English 2000) and 428 from the Brezina and Gablasova's (2015) New-General Service List (New-GSL). 78 English language teachers rated the usefulness of each of the headwords on a five-point Likert scale (1 being the least useful and 5 being the most useful) based on how these words would help their students perform basic functions in English. In parallel, 135 Vietnamese learners of English took *yes/no* tests that included the same 973 words, as well as 408 pseudowords, thus providing reliable information about students' reported knowledge of the words. Words from the BNC/COCA2000 —which relies both on corpus frequency and subjective usefulness ratings— were considered more useful by the teachers and were better known by the students than the words from the New-GSL, which only computes corpus-based frequency with no subjective ratings of usefulness. These data reinforce the idea that wordlists need to be developed through a balanced approach that takes into account objective measures of lexical frequency, student's lexical knowledge, and teachers' considerations about words' usefulness.

These considerations will be helpful in designing future wordlists and textbooks but the very real and immediate problem that Spanish language instructors face in their current classes is that, while such lists and textbooks do not yet exist, they still need to make daily decisions about which words to teach. As was stated above, said decisions could draw on (1) corpus-based word frequency data, (2) teachers' perception of usefulness, and (3) students' actual vocabulary knowledge or, (4) ideally, a combination of those three criteria. Concerning lexical frequency, studies have proved that teachers tend not to make active use of corpus-based frequency lists (Dang and Webb 2020) and have relatively poor intuitions about specific word frequencies beyond the 2,000 most frequent ones (McCrostie 2007). With respect to usefulness, it has mostly been conceptualized in terms of a words' relevance to fulfill functions that are specific to the L2 classroom (Dang, Webb, and Coxhead 2020; Garnier and Schmitt 2015; He and Godfroid 2019; Laufer and Nation 2012; Stein 2017). For instance, a word such as *adjective* may not be frequent in a corpus of native speakers, but it is definitely useful when first discovering the notion of gender and number agreement in a beginner Spanish course. However, there is an aspect of a word's usefulness that has yet to be explored and that connects points (2) and (3) of the list of

vocabulary selection criteria presented earlier: teachers' intuitions about L2 learners' lexical knowledge at different levels of proficiency. Indeed, a teacher will naturally find that it is more useful to spend time explicitly teaching and practicing a given word if they believe that most of their students do not already know it.

Coming back to the *zucchini* vs. *eggplant* conundrum, if a pure frequency criterion was used, only *eggplant* would be taught but, really, who would believe that not one student in the class would actually prefer to learn *zucchini* (or would want to learn both)? If, instead, a usefulness criterion was favored, there would be no way to make a reasonable decision, as a preference for *zucchini* over *eggplant*, or vice versa, would depend more on the recipes that students have in mind when learning how to buy food at the supermarket than on any inherent level of usefulness of either vegetable. In this example, the last criterion, namely students' actual knowledge of the word may seem like the most reasonable option for vocabulary selection. Since both words can be just as useful depending on each student's culinary goals in performing the *buying food at the supermarket* task, the word that would actually be more useful to focus on explicitly in class would be the one that students do not yet know.

Although this criterion may seem sound, if no systematic vocabulary tests are completed by the students to provide evidence of their learning process, it completely relies on teachers' intuitions about their students' lexical knowledge. And inaccurate intuitions may result in inefficient teaching decisions, where some words are dismissed due to the erroneous belief that students know them, and conversely others that are already mastered are taught over and over again. If these types of errors were to happen often in a same language course, students could feel overwhelmed while flooded with unknown words or, alternatively, bored and unchallenged in a class where no new words are introduced. The present study aims to explore whether teachers are good judges of students' lexical knowledge or not by addressing these specific research questions (RQs):

RQ1: How accurate are teachers' intuitions about L2 students' reported lexical knowledge?

RQ2: Are teachers' intuitions more, or less, accurate at different course levels?

RQ3: In course levels where teachers tend to have less accurate judgments of their students' reported lexical knowledge, are teachers' intuitions more, or less, accurate for words in different frequency ranks?

RQ4: Does the amount of teaching experience at a particular course level impact the accuracy of teachers' judgment accuracy in that level?

2 Methods

2.1. Participants

Participants in this study came from two pools: (1) 421 students enrolled in one of the three courses of a First-Year Spanish program at a large public West Coast University (i.e., SPA 1, 2 and 3), and (2) 38 instructors who taught those courses at the same institution. The First-Year Spanish program is designed for students who start learning Spanish with no previous

experience with the language. Given that there was only one instructor per each class of approximately 25 students, and that those instructors were the same across academic terms (i.e., 11-weeks long quarters), the number of participants in the student pool is significantly larger than that in the instructors' pool. Nevertheless, those instructors represent 100% of the teacher population for those courses during the time of data collection.

From the 421 participants in the student pool, 216 were enrolled in SPA 1 (i.e., the first course in the program), 138 in SPA 2 (i.e., the second course in the series), and 117 in SPA 3 (i.e., the third course in the program). Of those, 70.2% were female, 18.1% male and 1.7% selected the "Other" option when asked about their gender identity. The mean age of the student participants was 20.32, ranging from a minimum of 17 to a maximum of 58. Concerning their L1s, 80.1% declared that English was their first language, 11.6% chose Mandarin and the other 8.3% had a variety of other L1s, such as Vietnamese, Hmong or Japanese.

When it comes to the 38 instructors who participated in this study, 13 were teaching SPA 1, 10 were teaching SPA 2, and 15 were teaching SPA 3. 68.4% of them were female and 28.9% were male. One instructor did not respond to the question about gender identity, which explains the missing 2.6%. The mean age of the instructors was 31.03, ranging from a minimum of 22 to a maximum of 51. 65.8% of the instructors reported that Spanish was their first language, whereas 28.9% declared that it was English. The remaining 5.3% (N=2) wrote another language as their L1. On average, they had taught the course they were offering at the time 2.18 times, with a minimum of 1 and a maximum of 6.

2.2. Items

Both students and instructors were presented the full list of real words included in 3K-LEx[1], (Robles-García 2020a), a *yes*/*no* lexical decision test in which participants indicate whether or not they know the meaning of these words by saying *yes* or *no*, respectively. The complete test contains 108 real words (36 words pertaining to each of the three first 1,000-frequency bands) which were randomly selected from Davies and Davies (2017) Spanish lemmatized frequency dictionary, as well as 54 pseudowords, which were not used in this study. 3K-LEx measures the written recognition (Nation 2013) of nouns, verbs, and adjectives. Reflecting the actual percentage of these word classes within the 3,000 most frequent words in Spanish, 3K-LEx follows a 60 (noun): 28 (verb): 20 (adjective) ratio.

2.3. Data collection

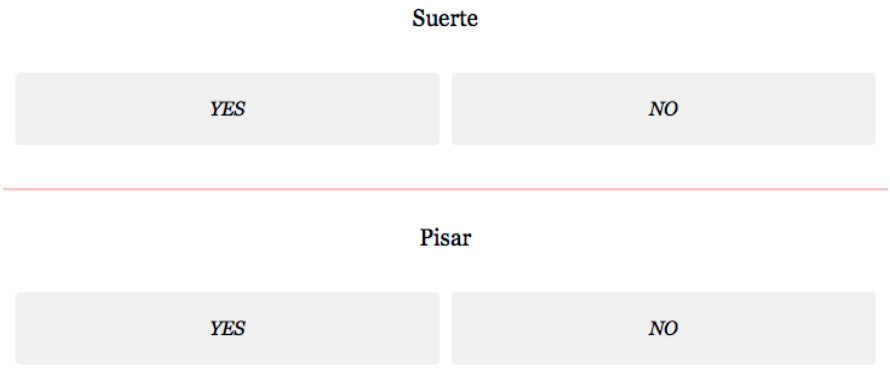
The test for the students was administered during the last two weeks of the course to ensure highest levels of vocabulary knowledge. Likewise, instructors were also asked to take the test during the same period of time. Data collection took place during the three academic terms of the academic year 2018-2019. Both groups of participants completed the corresponding tests on Qualtrics (2002), but students did so during scheduled class time while teachers could

[1] 3K-LEx can be downloaded for free from the following link: https://github.com/problg00/3K-LEx. While the full 3K-LEx also includes a set of pseudowords, destined to be used in a penalization matrix to calculate the total scores in the test, this study only uses data from the 108 real words in the test.

complete the questionnaire on their own time. The following image illustrates the format of two items from the test:

Figure 1

Example of 3K-Lex Test Items



Before administering the test, all participants responded to a short demographic questionnaire and signed a consent form in which information about the nature of the test and their rights as participants were explained in detail. Students were again orally reassured that their data would be de-identified upon their receiving the corresponding extra credit that was approved on the IRB protocol and that their instructors would not have access to their results.

Students were told that they would see a series of words on the screen, and they would have to decide whether or not they knew the meaning of each of them by pressing *yes*, if they did, or *no*, if they did not. They were encouraged to select *no* whenever they were unsure about the meaning of any given word and to limit their *yes* responses to words for which they would easily give a translation or a definition. Students were not given any time restriction in completing the test —other than finishing during class time, which amounted to 50 minutes. They all finished within 12 minutes.

Instructors were told that they would see a series of words and would have to decide whether or not they thought students at the course level they were teaching at the time of testing knew the meaning of the word. If they thought most students did know the word, they were asked to select *yes*, and if they thought they did not, they were asked to respond *no*. If they were unsure, we asked them to select *no*, and to reserve their *yes* responses strictly for words that they were convinced most students in their course level would know. Since teachers were explained how the students were taking the test (i.e., responding *yes* when they recognized the word, and *no* when they did not), they were asked to think about word knowledge in terms of word recognition, which was the type of knowledge assessed in the students' test. They were allowed to complete the test any time within the last two weeks of the academic term, and all of them took less than 17 minutes to complete the test.

2.4. Measurements

A total of 108 scores was obtained for the student dataset, and another 108 for the teacher dataset, one score per word in the test (see Appendix 1). Each score represents the percentage of *yes* responses per item[2]. Concretely, a score of 0 for a given item means that 0% of the participants responded *yes* to this item, whereas a score of 1 indicates that 100% of the participants responded *yes*. For example, a score of 0.6 in the students' dataset for a given item would mean that 60% of the students who participated declared knowing the word. A score of 0.4 for that same item in the teachers' dataset would mean that 40% of the teachers thought that the students enrolled in the same course level they were teaching would know that word. In sum, the unit of analysis for each item is the percentage of participants in each dataset (i.e., students vs. teachers) who responded *yes* to that particular item. Exploratory comparisons between groups were thus carried out by item and not by subject, given the great difference in sample size between groups.

It is important to note that since responses were anonymous in both datasets (i.e., students and teachers), it was impossible to match a teacher's response and the responses of the students enrolled in their own class at the times of testing.

3 Results

Since the data in both groups were not normally distributed, as assessed through Kolmogorov-Smirnov tests, and were skewed towards 0 (i.e., there were more words that students declared to not know than to know, and that teachers thought their students did not know than know), only non-parametric analyses were carried out in this study.

Results are presented in three steps. First, teachers' and students' results were assessed through a correlational analysis. This initial step aimed to respond to RQ1. The second step examined students' and teachers' responses further, through a series of non-parametric tests and visual examinations of response distribution graphs. This step aimed to answer RQ2 and RQ3. Finally, another correlational analysis was carried out, which explored the relationship between teachers' experience and the accuracy of their intuitions regarding students' reported vocabulary knowledge, thus responding to RQ4.

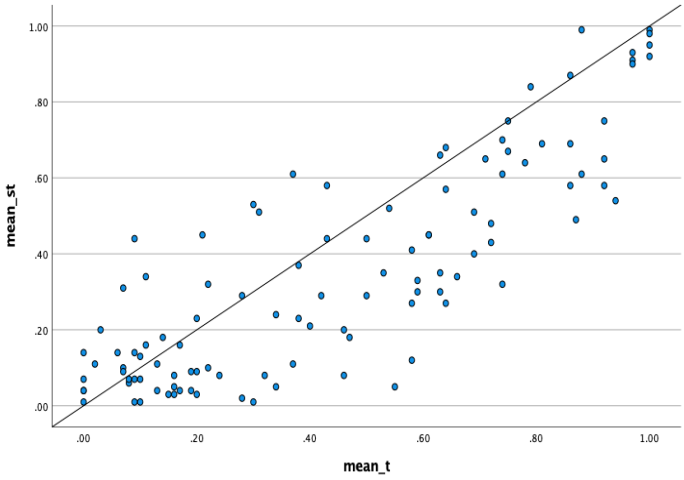
3.1. RQ1: How accurate are teachers' intuitions about L2 students' reported lexical knowledge?

One Spearman correlation was run using (1) the percentage of *yes* responses per item across the 421 student participants, and (2) the percentage of *yes* responses per item across the 38 teacher participants. Results indicate that the correlation was very high and significant, r=.811, p <.001, which is also visually confirmed in Figure 2.

Figure 2

[2] Even though our data are not normally distributed since there are many more words that beginner students do not know than words that they know, resulting in a high kurtosis and data that are skewed left, we did not carry out the analyses utilizing median values. The median of series of 1 and 0 responses results in a median of either 1 or 0 per item, thus not providing enough distribution to carry out any statistical analysis. Instead, the percentage of *yes* responses is a measure that, although imperfect, makes results easy to interpret.

Scatter Plot of Percentage of Yes Responses in Teachers (Mean_t) and Mean Percentage of Yes Responses in Student Responses (Mean_st)



Even though data from students and teachers seem to be highly correlated, it is important to know whether student' and teachers' responses show similar distributions across course levels. This will be further explored in the next section, where the distribution of teachers' and students' responses will be compared at each course level.

3.2. *RQ2: Are teachers' intuitions more, or less, accurate at different course levels?*

Three Mann-Whitney U tests were run, one per each course level, to compare students' and teachers' percentage of *yes* responses to each of the words in the test. Table 1 presents the summary of results for each of the tests and reveals that the distribution of percentages of *yes* responses between groups is only significantly different in SPA 2, with an almost medium effect size of 0.23. This result is further confirmed through the visual exploration of Figure 3, which shows that while students' data at that level are skewed towards 0 (i.e., not knowing the word), teachers' data present the opposite pattern, being skewed towards 1 (i.e., knowing the word).

Table 1

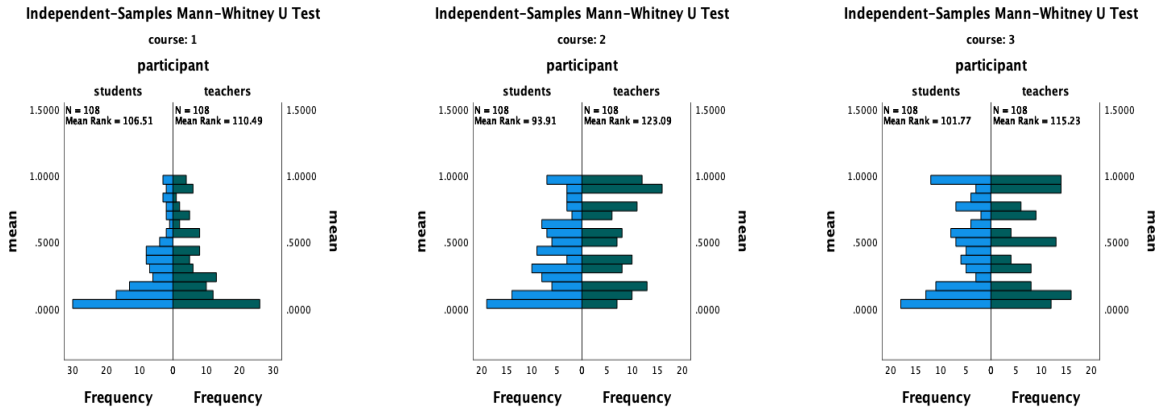
Independent-Samples Mann-Whitney U Tests of responses by Course Level

	SPA 1	SPA 2	SPA 3
Mann-Whitney U	6047.000	7407.500	6558.500
Wilcoxon W	11933.000	13293.500	12444.500

Standard Error	458.658	458.914	459.040
Asymptotic significance (2-sided test)	0.639	0.001*	0.114
Effect size (*r*)	0.11	0.23	0.03

Figure 3

Histograms of the Distribution of Responses by Teachers and Students in SPA 1 (course 1), SPA 2 (course 2) and SPA 3 (course 3)



Since teachers in SPA 2 are the ones who present the distribution of responses that differs the most from that of the students enrolled at that same course level, it would be interesting to verify if there are words at specific frequency ranks that are driving those differences. Concretely, are teachers in SPA 2 less accurate in assessing students' knowledge of words at lower frequency ranks?

3.3. *RQ3: In course levels where teachers tend to have less accurate judgments of their students' lexical knowledge, are teachers' intuitions more, or less, accurate for words in different frequency ranks?*

Three Mann Whitney U tests were run, one per word frequency rank (i.e., rank 1=1-1,000 most frequent words, rank 2 = 1,001-2,000 most frequent words, rank 3 = 2,001-3,000 most frequent words), comparing teachers' and students' percentages of *yes* responses in SPA 2. Results indicate that significant differences arose at all frequency ranks with medium (or close to medium) effect sizes, showing that teachers tended to overestimate their students' knowledge at all frequency ranks, and not only for lower frequency words.

Table 2
Independent-Samples Mann-Whitney U Tests by Frequency Ranks in SPA 2

	Rank 1	Rank 2	Rank 3
Mann-Whitney U	892.000	895.500	834.000
Wilcoxon W	1558.000	1561.500	1500.000
Standard Error	88.402	88.681	88.633
Asymptotic significance (2-sided test)	0.006*	0.005*	0.036*
Effect size (*r*)	0.32	0.32	0.24

3.4. *RQ4: Does the amount of teaching experience at a particular course level impact the accuracy of teachers' judgment accuracy in that level?*

This study also aims to research the effect of an instructor's teaching experience at a particular course level on their accuracy levels when predicting students' knowledge (as reported by the students themselves) of the words in the test. In order to obtain a parameter that could approximate an instructor's closeness to the real knowledge of students enrolled in the course level they teach, we subtracted the mean result of the students at that level from each teacher's average response. For example, in SPA 1, the average knowledge of the students was of 0.26 (i.e. 26%). If an instructor's average was 0.56 (i.e., 56%), the difference would be of 0.3 (i.e., 30%). Upon making this calculation for each instructor, a Spearman correlation was run between this difference and the number of times that the instructor had taught that particular course level. Results from this analysis did not show any significant correlation (r= -.185; p= .267), indicating that the teachers' distance to students' declared knowledge were not correlated with their amount of experience teaching that particular course level. This analysis, however, is a first approximation to this question and would need to be re-run with a larger sample of teachers in future studies.

4 Discussion

The goal of this study was to research Spanish teachers' intuitions about their students' lexical knowledge by addressing the following questions:

RQ1: How accurate are teachers' intuitions about L2 students' reported lexical knowledge?

RQ2: Are teachers' intuitions more, or less, accurate at different course levels?

RQ3: In course levels where teachers tend to have less accurate judgments of their students' reported lexical knowledge, are teachers' intuitions more, or less, accurate for words in different frequency ranks?

RQ4: Does the amount of teaching experience at a particular course level impact the accuracy of teachers' judgment accuracy in that level?

With respect to RQ1, the quick response is: *Yes*! Instructors in this study tended to be overall accurate in their intuitions about the lexical knowledge of students enrolled in the course level they were teaching at the time, as revealed in the strong correlation between students' and teachers' responses. However, when comparing the distribution of *yes* responses in both groups of participants at each course level, teachers' responses tended to present very similar distributions to those of students in SPA 1 and SPA 3, while SPA 2 teachers tended to believe that many words were known by the students when students had declared not knowing them. Therefore, in response to RQ2, the accuracy of teachers' intuition does seem to depend on the course level, with overall lower accuracy in SPA 2. Given the teaching context (i.e., three consecutive courses taught by TAs who generally start teaching SPA 1 and move on to SPA 2 and 3), these results may be due to the fact that instructors in SPA 3 had generally taught SPA 1 and 2 in the past, which provided them an advantageous longitudinal overview of all the levels. Indeed, not only did they have experience with each of the chapters of the book used in all three courses, which may have provided useful information about the vocabulary that students are exposed to (Cubillos 2014), but they also had personally interacted with students at all levels by teaching all courses in the language program. Teachers in SPA 1, on the other extreme, only had experience working with SPA 1 students and (accurately) assumed that these knew very few words. Instructors in SPA 2 found themselves in an uncomfortable in-between position, where they probably realized that their current students knew more words than those in SPA 1 but they had not yet been exposed to SPA 3 students. This incomplete set of information may have favored a tendency towards overestimation. Moreover, their misguided intuition may also have been driven by the very real fact that the percentage of additional words learned from SPA 1 to SPA 2 is much higher, according to this study's data, (i.e., 12%) than that observed between students in SPA 2 and those in SPA 3 (i.e., 4%).

In response to RQ3, when digging deeper into the specific words for which SPA 2 teachers tended to overestimate students' knowledge, no differences arose between frequency ranks. Concretely, teachers overestimated their students' knowledge for words that were among the first most frequent in Spanish, but also for lower frequency words. It may be that teachers do not rely as much on a frequency criterion when determining whether a word is expected to be known or not known by students at a particular course level, or that, in line with what was found in previous research, they may not have a fine-tuned awareness of frequency differences between words (McCrostie 2007). What we do know from previous research is that teachers give a lot of

importance to a words' usefulness (Dang, Webb, and Coxhead 2020; Garnier and Schmitt 2015; He and Godfroid 2019; Laufer and Nation 2012; Stein 2017), which can provide a potential explanation for some of the words for which teachers tended to overestimate students declared knowledge. For instance, in SPA 1, 54% of instructors thought that their students knew the verb *borrar* 'to erase', which is in frequency rank 3, whereas only 16% of students declared knowing it. This word, although not highly frequent in native speakers' Spanish corpora, is used repeatedly in the classroom when erasing information from the blackboard. Similar trends are observed with other words that are used often in the classroom but are not very frequent in a native corpus, such as *adivinar* 'to guess' or *desafío* 'challenge', which teachers may generally think are known because they are regularly utilized when giving instructions about typical exercises and activities.

Another source of information about what students may or may not know is the presence or absence of words in textbook activities and glossaries. As is discussed in Marcos Miguel and Cubas in this volume, teachers greatly rely on textbook contents and activities when teaching vocabulary in their classes. In this context, it would not be surprising if they (sometimes wrongly) assumed that words that are included in the textbook are known by the students. An interesting example of one such word is *alimento* 'food item', which 90% of the teachers in SPA 2 thought was mastered by their students when only 27% of them reported knowing it. The word is in frequency rank 2 but generally serves as a title for lists of food items in textbook glossaries and it would be easy to imagine that teachers use it when talking about food in general. However, it does not point to any specific food item that students can use when going grocery shopping or ordering a dish in a restaurant. This could be an example of how teachers' intuitions may be partially biased by what they know their students are exposed to in the textbooks, rather than on the observation of actual students' use, or recognition, of the words. However, research shows that simple exposure to a word does not ensure incidental learning unless occurrences of a word are frequent enough (i.e., 10 to 20 repetitions) and that words that are taught explicitly in class have a better chance of being learned (Dóczi and Kormos 2016; Folse 2010). Thus, even though students may have seen the word *alimento* in the textbook, if no particular attention was given to it since it was not necessary to complete any of the activities, there would be little chances of them learning the word.

This last idea calls the attention to the fact that, when we talk about lexical frequency, we are referring to calculations that are based on native speaker corpora. If we want to better understand the impact of lexical frequency on students' word recognition and lexical knowledge, frequency counts need to be extracted from corpora of classroom recordings. Even if textbooks can serve as "proxies" of what happens in the classroom, they may not be good indicators of how much emphasis is given to specific words. For example, knowing that the word *alimento* appears in the textbook is not enough to assume that it is repeated enough times during class to facilitate incidental learning, or that it is taught explicitly. Having the information about what ultimately happens in the classroom would allow us to use frequency counts that correspond to the actual input students are exposed to and to better understand the relation between what appears in the textbooks and what teachers do with it. This information, in turn, would make it possible to assess whether students' and teachers' responses in studies similar to ours may be driven by

classroom frequency and explicit treatment of certain words, instead of relying on native speakers' frequency counts which are disconnected from the realities of L2 classrooms. Unfortunately, no such corpus has been created and shared so far, to the best of our knowledge.

The hypotheses proposed above about the driving forces behind teachers' intuitions of student learning currently remain at the hypothetical level, but they would certainly deserve further inquiry in future studies on L2 teachers' intuitions about students' lexical knowledge. Importantly, in response to RQ4, no matter the criteria that may have driven teachers' responses, the accuracy of said responses does not seem to be influenced by the level of experience that the teacher has at that specific course level. However, again, this needs to be studied further, with more course levels and course types, as well as additional instructors with a broader range of teaching experiences.

Going back to the issue of vocabulary selection in L2 classes, the literature proposes three distinct criteria that could prove useful in fulfilling this arduous task: (1) corpus-based word frequency data (Schmitt and Schmitt 2014), (2) teachers' perception of usefulness (Dang, Webb, and Coxhead 2020; Garnier and Schmitt 2015; He and Godfroid 2019; Laufer and Nation 2012), and (3) students' actual vocabulary knowledge (Brysbaert, Keuleers, and Mandera 2020). Based on previous research, teachers do not rely on the first source of data (Dang and Webb 2020; Sánchez-Gutiérrez, Robles-García, and Pérez Serrano 2022), and very few currently available wordlists incorporate teachers' subjective ratings of usefulness (Dang, Webb, and Coxhead 2020). In the case of Spanish, there is simply none of the sort. To date, there are no wordlists in Spanish that adopt the bottom-up approach proposed by Brysbaert, Keuleers, and Mandera (2020), which entails organizing words based on the number of students who know them. Because no such database exists in Spanish, teachers have to rely on their own intuitions about which words their students may know and which they may not in order to make daily decisions about vocabulary teaching. As evidenced in this study, teachers have relatively accurate perceptions of their students' reported word knowledge, especially when those are beginners or when they are taking the last course in a first-year language program. Nevertheless, there are still mismatches between students' reported knowledge and the knowledge that is assumed from them by their teachers. This is particularly true for certain specific words that teachers may want their students to know and thus assume most of their students acquired. While more research is needed in this respect, those inconsistencies in teachers' and students' data demonstrate that vocabulary selection decisions should not rely solely on teacher's intuitions but rather on observed student knowledge, assessed by consistent vocabulary tests. Therefore, we propose that teachers increase the number and frequency of vocabulary assessments in their courses in order to gather reliable data about their students' lexicon, which will in turn allow them to make better decisions about which words to teach.

5 Conclusion

This study aimed to explore the extent to which L2 Spanish teachers are good judges of their students' reported lexical knowledge. Upon analyzing the results of a self-reported

vocabulary test that was completed by students enrolled in three consecutive beginner language courses and comparing them to the assessment of their word knowledge provided by teachers in those same courses, teachers seem to have a relatively accurate perception of students' reported knowledge. However, this is far from being true with all words in the test, and more research is needed that analyzes the criteria that underlie teachers' intuitions about their students' (self-reported) knowledge of certain vocabulary. While the creation of a wordlist that determines the words that should be taught at different proficiency levels based on a large-scale inquiry of student data would be much welcome in the longer term, teachers can, in the meantime, systematically assess their students' lexical knowledge through continuous testing.

While this study offers an initial exploration into teachers' awareness about students' vocabulary knowledge, it presents several limitations. First, *yes/no* tests, such as the one used in this study, are known to overestimate learners' knowledge, since they do not require to provide any evidence of actually knowing the meaning of the word and thus depend on learners' reported knowledge rather than proof of actual knowledge (Schmitt et al. 2020; Stoeckel et al. 2021). However, at the moment, no meaning recall vocabulary test has been designed in Spanish that is validated and published for public use, which limits researchers' ability to investigate students' actual lexical knowledge. Another limitation of the test is that it focuses only on one type of vocabulary knowledge: the form-meaning link. As is pointed out in Zyzik & Marqués Pascual (this volume), the overfocus on this aspect of lexical knowledge does not reveal the whole picture of what knowing a word means. For instance, it may be that students would not be able to translate a word directly but would know that it is related to other words in its lexical family, thus revealing knowledge of word parts. In future studies, as more validated vocabulary tests become available in Spanish, it would be advisable to use a meaning-recall test, but it would be even better to test different aspects of word knowledge as well.

Additionally, teachers in this study were asked to respond to the test items thinking about the probability that a general group of students (e.g., those in SPA 3) would know a word. Given that students in a class may vary greatly in the amount of vocabulary they know, teachers may provide responses that either focus on the students they think know the most or the ones that know the least. Depending on what subgroup of students they were considering when completing the test, they may have over-or under-estimated their knowledge, adding a lot of variability to the teachers' results. In future studies, it would be better to match teachers with students from the specific class they are teaching at the time of testing. An alternative (or perhaps complementary) way of addressing this issue would be to provide teachers with a scale instead of a *yes/no* response, which would allow them to indicate how sure they are that students at that proficiency level would know each word or not or what percentage of their students they think would know it. Finally, this study was merely exploratory and only looked at beginner courses and at a limited number of instructors. Future studies would benefit from adding more proficiency levels and including a larger number of instructors.

References

Aizawa, Kazumi, Masamichi Mochizuki & Paul Meara. 2001. Intuition about word frequency: What does it tell us about vocabulary knowledge. *Research Reports of the Faculty of Engineering* 20. 75–82. Tokyo Denki University.

Alcaraz Mármol, Gema. 2009. Vocabulary in EFL textbooks: Frequency levels. In Pascual Cantos Gómez & Aquilino Sánchez Pérez (eds.), *A survey on corpus-based studies*, 862–875. http://www.um.es/lacell/aelinco/contenido/index.html (accessed 12 March 2021).

Alderson, Charles. 2007. Judging the frequency of English words. *Applied Linguistics* 28 (3). 383–409. https://doi.org/10.1093/applin/amm024

Brezina, Vaclav & Dana Gablasova. 2015. Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics* 36. 1–22. https://doi.org/10.1093/applin/amt018

Brysbaert, Marc & Boris New. 2009. Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41. 977–990. https://doi.org/10.3758/BRM.41.4.977

Brysbaert, Marc, Emmanuel Keuleers & Paweł Mandera. 2020. Which words do English non-native speakers know? New supernational levels based on yes/no decision. *Second Language Research* 37 (2). 207–231. https://doi.org/10.1177/0267658320934526

Creighton, Theodore B. 2007. *School and data: The educator's guide for using data to improve decision making*. London: Sage.

Criado, Raquel & Aquilino Sánchez. 2009. Vocabulary in EFL textbooks: A contrastive analysis against three corpus-based word ranges. In Pascual Cantos Gómez & Aquilino Sánchez Pérez (eds.), *A survey on corpus-based studies*, 862–875. http://www.um.es/lacell/aelinco/contenido/index.html (accessed 12 March 2021).

Cubillos, Jorge H. 2014. Spanish textbooks in the US: Enduring traditions and emerging trends. *Journal of Spanish Language Teaching* 1 (2). 205–225. https://doi.org/10.1080/23247797.2014.970363

Daller, Michael, John Turlik & Ian Weir. 2013. Modelling vocabulary growth in a foreign language: Defining and measuring lexical diversity. In Scott Jarvis & Michael Daller (eds.), *Vocabulary knowledge: Human ratings and automated measures*, 185–218. Amsterdam: John Benjamins.

Dang, Thi Ngoc Yen & Stuart Webb. 2020. Vocabulary and good language teachers. In Carol Griffiths & Zia Tajeddin (eds.), *Lessons from good language teachers*, 203–218. Cambridge University Press.

Dang, Thi Ngoc Yen, Stuart Webb & Averil Coxhead. 2020. Evaluating lists of high-frequency words: Teachers' and learners' perspectives. *Language Teaching Research*, Online First, 1–25. https://doi.org/10.1177/1362168820911189

Davies, Mark. 2005. Vocabulary range and text coverage: Insights from the forthcoming Routledge frequency dictionary of Spanish. In David Eddington (ed.), *Selected proceedings of the 7th Hispanic Linguistics Symposium*, 106–115. Somerville, MA: Cascadilla Proceedings Project.

Davies, Mark, & Kathy Hayward Davies. 2017. *A frequency dictionary of Spanish: Core vocabulary for learners, Second Edition.* New York: Routledge.

Davies, Mark & Timothy L. Face. 2006. Vocabulary coverage in Spanish textbooks: How representative is it? In Nuria Sagarra & Almeida Jacqueline Toribio (eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, 132–143. Somerville, MA: Cascadilla.

Donzelli, Giovanna. 2007. Foreign language learners: Words they hear and words they learn: A case study. *Estudios de Lingüística Inglesa Aplicada* 7. 103–125. https://revistas.uned.es/index.php/ELIA/article/view/18092

Dóczi, Brigitta & Judit Kormos. 2016. *Longitudinal developments in vocabulary knowledge and lexical organization*. Oxford: Oxford University Press.

Earl, Lorna & Steven Katz. 2006. *Leading in a data rich world*. Thousand Oaks, California: Corwin press.

Epstein, Seymour. 2008. *Intuition from the perspective of cognitive experiential self-theory*. New York: Erlbaum.

Folse, Keith. 2010. Is explicit vocabulary focus the reading teacher's job? *Reading in a Foreign Language* 22 (1). 139–160. http://nflrc.hawaii.edu/rfl

Garnier, Mélodie & Norbert Schmitt. 2015. The PHaVE List: A pedagogical list of phrasal verbs and their most frequent meaning senses. *Language Teaching Research* 19 (6). 645–666. https://doi.org/10.1177/1362168814559798

He, Xuehong & Aline Godfroid. 2019. Choosing words to teach: A novel method for vocabulary selection and its practical application. *TESOL Quarterly* 53 (2). 348–371. https://doi.org/10.1002/tesq.483

Horst, Marlise. 2013. Mainstreaming second language vocabulary acquisition. *The Canadian Journal of Applied Linguistics* 16 (1). 171–188. https://journals.lib.unb.ca/index.php/CJAL/article/view/21299

Laufer, Batia. 1989. What percentage of text-lexis is essential for comprehension? In Christer Lauren & Marianne Nordman (eds.), *Special language: From humans thinking to thinking* machines, 316–323. Clevedon: Multilingual Matters.

Laufer, Batia & Paul Nation. 1999. Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics* 16 (3). 307–322. https://doi.org/10.1093/applin/16.3.307

Laufer, Batia & Geke C. Ravenhorst-Kalovski. 2010. Lexical threshold revisited: Lexical coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language* 22 (1). 15–30. http://nflrc.hawaii.edu/rfl

Laufer, Batia & Paul Nation. 2012. Vocabulary. In Susan M. Gass & Alison Mackey (eds.), *The Routledge Handbook of Second Language Acquisition*, 163–176. London, England: Routledge.

Lipinski, Silke. 2010. A frequency analysis of vocabulary in three first-year textbooks of German. *Die Unterrichtspraxis/Teaching German* 43 (2). 167–174. https://doi.org/10.1111/j.1756-1221.2010.00078.x

López Bastidas, Lani & Claudia Sánchez-Gutiérrez. 2020. Vocabulary selection and word repetitions in beginner L2 Spanish textbooks. *Revista Electrónica de Lingüística Aplicada* 19 (2). 48-63. http//www.aesla.org.es/ojs/index.php/RAEL/article/view/432

Martini Oliveira Pisani, Juliane. 2012. *High frequency vocabulary in a secondary Quebec ESL textbook corpus*. Montreal, Canada: Concordia University MA thesis. https://spectrum.library.concordia.ca/id/eprint/974698/

McCrostie, James. 2005. A further investigation of word frequency intuitions. *The East Asian Learner* 2 (1). 26–36. https://doi.org/10.1177/0033688206076158

McCrostie, James. 2007. Investigating the Accuracy of Teachers' Word Frequency Intuitions. *RELC Journal* 38 (1). 53–66. https://doi.org/10.1177/0033688206076158

Meara, Paul. 1980. Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching and Linguistics Abstracts* 13 (3-4). 221–246. https://doi.org/10.1017/S0261444800008879

Milton, James. 2009. *Measuring second language vocabulary acquisition*. Bristol, UK: Multilingual Matters.

Nation, Paul. 2006. How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review* 63 (1). 59–82. https:/7doi.org/10.3138/cmlr.63.1.59

Nation, Paul. 2012. *The BNC/COCA word family lists*. http://www.victoria.ac.nz/lals/about/staff/paul-nation (accessed 20 March 2021).

Nation, Paul. 2013. *Learning vocabulary in another language* (2nd edn.). Cambridge: Cambridge University Press.

Ringeling, Tobi. 1984. Subjective Estimations as a Useful Alternative to Word Frequency Counts. *Interlanguage Studies Bulletin* 8. 59–69. https://www.jstor.org/stable/43135301

Robles-García, Pablo. 2020a. 3K-LEx. Desarrollo y validación de una prueba de amplitud léxica en español. *Journal of Spanish Language Teaching* 7 (1). 1–13.

Robles-García, Pablo. 2020b. *¿Cuántas palabras conocen los aprendientes de español como lengua extranjera? Competencia léxica y oportunidades de aprendizaje léxico en el aula de ELE.* Davis, CA: University of California Davis dissertation.

Sánchez-Gutiérrez, Claudia H., Pablo Robles-García & Mercedes Pérez Serrano. 2022. L2 Spanish vocabulary teaching in US universities: Instructors' beliefs and reported practices. *Language Teaching Research*. Online First. 1–23.

Sánchez-Gutiérrez, Claudia, Nausica Marcos Miguel & Michael K. Olsen. 2018. Vocabulary coverage and lexical characteristics in L2 Spanish textbooks. In Peter Ecke & Susanne Rott (eds.), *Understanding vocabulary learning and teaching: Implications for language program development*, 78–98. Boston, MA: Cengage. http://hdl.handle.net/10125/69783

Schmitt, Norbert & Bruce Dunham. 1999. Exploring native and non-native intuitions of word Frequency. *Second Language Research* 15 (4). 389–411. https://doi.org/10.1191/026765899669633186

Schmitt, Norbert & Diane Schmitt. 2014. A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching* 47 (4). 484–503. https://doi.org/10.1017/S0261444812000018

Schmitt, Norbert, Karen Dunn, Barry O'Sullivan, Laurence Anthony & Benjamin Kremmel. 2020. Introducing knowledge-based vocabulary lists (KVL). *TESOL Journal*, 12 (4). 1–10. https://doi.org/10.1002/tesj.622

Stæhr, Lars Stenius. 2008. Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal* 36. 139–152. https://doi.org/10.1080/09571730802389975

Stein, Gabriele. 2017. Some thoughts on the issue of core vocabularies: A response to Vaclav Brezina and Dana Gablasova: 'Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics* 38 (5). 759–763. https://doi.org/10.1093/applin/amw027

Stoeckel, Tim, Stuart McLean & Paul Nation. 2021. Limitations of size and levels tests of written receptive vocabulary knowledge. *Studies in Second Language Acquisition* 43 (1). 181–203. https://doi.org/10.1017/S027226312000025X

Sun, Ye & Thi Ngoc Yen Dang. 2020. Vocabulary in high-school EFL textbooks: Texts and learner knowledge. *System* 93. 1–13. https://doi.org/10.1016/j.system.2020.102279

Vanlommel, Kristin, Ross Van Gasse, Jan Vanhoof & Peter Van Petegem. 2017. Teachers' decision-making: Data based or intuition driven? *International Journal of Educational Research* 83. 75–83. https://doi.org/10.1016/j.ijer.2017.02.013

Van Zeeland, Hilde & Norbert Schmitt. 2013. Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics* 34 (4). 457–479. https://doi.org/10.1093/applin/ams074

Yang, Lu & Averil Coxhead. 2020. A corpus-based study of vocabulary in the New Concept English textbook series. *RELC Journal*. Online First. 1–15. https://doi.org/10.1177/0033688220964162

Appendix 1

	SPA 1		SPA 2		SPA 3	
item	students	teachers	students	teachers	students	teachers
encontrar	0.58	0.85	0.85	0.90	0.81	1.00
pequeño	0.83	0.92	0.95	1.00	0.96	1.00
esperar	0.40	0.54	0.72	0.90	0.96	1.00
tiempo	0.88	1.00	0.99	1.00	0.99	1.00
trabajo	0.98	1.00	1.00	1.00	0.98	1.00
solo	0.76	0.69	0.91	0.90	0.95	1.00
cuerpo	0.28	0.77	0.56	1.00	0.89	1.00
creer	0.68	0.46	0.87	0.90	0.96	1.00
mano	0.80	0.92	0.90	1.00	0.99	1.00
todo	0.86	1.00	0.94	1.00	0.95	1.00
meter	0.35	0.00	0.52	0.20	0.44	0.07
cara	0.39	0.69	0.65	1.00	0.80	0.93
mostrar	0.14	0.23	0.28	0.80	0.49	0.87
empresa	0.41	0.54	0.55	0.80	0.56	0.73
fuerza	0.11	0.31	0.44	0.80	0.36	0.67
dinero	0.82	0.92	0.97	1.00	1.00	1.00
descubrir	0.68	0.54	0.58	0.80	0.70	0.80
antiguo	0.23	0.23	0.50	0.80	0.62	0.80
lanzar	0.04	0.00	0.06	0.10	0.09	0.13
suerte	0.50	0.69	0.68	0.70	0.82	0.87
asunto	0.03	0.08	0.05	0.10	0.03	0.20
dolor	0.35	0.38	0.60	0.80	0.76	0.73
crecer	0.32	0.54	0.42	0.50	0.56	0.47
cerrar	0.40	0.69	0.60	0.90	0.74	1.00
enviar	0.24	0.46	0.62	0.90	0.58	0.80
rey	0.21	0.46	0.33	1.00	0.47	0.53
abierto	0.23	0.77	0.58	0.90	0.65	0.93
vivo	0.99	0.69	1.00	1.00	0.98	0.93
rato	0.35	0.23	0.65	0.40	0.59	0.27
lluvia	0.36	0.92	0.61	0.90	0.66	1.00
cuento	0.53	0.46	0.81	0.90	0.75	0.87
colocar	0.02	0.15	0.05	0.30	0.03	0.13
recuerdo	0.41	0.31	0.75	0.80	0.88	0.80
caballo	0.43	0.46	0.74	1.00	0.74	0.87
recoger	0.42	0.23	0.66	0.60	0.74	0.27
piedra	0.18	0.08	0.25	0.40	0.28	0.53

seco	0.11	0.08	0.24	0.60	0.33	0.47
esconder	0.10	0.15	0.26	0.80	0.18	0.47
vuelta	0.10	0.00	0.40	0.40	0.46	0.27
grito	0.20	0.00	0.44	0.60	0.46	0.53
encargar	0.08	0.00	0.19	0.30	0.14	0.00
ruido	0.19	0.38	0.25	0.90	0.61	0.60
labio	0.25	0.23	0.21	0.70	0.35	0.80
saltar	0.41	0.62	0.42	0.60	0.38	0.53
alimento	0.13	0.54	0.27	0.90	0.42	0.47
desnudo	0.06	0.00	0.12	0.40	0.06	0.33
temor	0.04	0.15	0.06	0.40	0.14	0.40
pérdida	0.12	0.00	0.32	0.20	0.26	0.40
confianza	0.18	0.31	0.32	0.60	0.36	0.60
cadena	0.05	0.00	0.08	0.10	0.08	0.20
bailar	0.96	1.00	1.00	1.00	0.99	1.00
cuello	0.30	0.38	0.24	0.90	0.74	0.87
robar	0.13	0.15	0.41	0.90	0.51	0.53
muestra	0.27	0.15	0.53	0.20	0.54	0.27
callar	0.12	0.31	0.14	0.40	0.07	0.40
esquina	0.11	0.23	0.41	0.40	0.36	0.20
lejano	0.01	0.23	0.02	0.40	0.01	0.27
fiel	0.06	0.15	0.09	0.10	0.18	0.13
carretera	0.10	0.31	0.30	0.60	0.21	0.47
traje	0.29	0.23	0.83	0.80	0.87	0.87
mentira	0.13	0.38	0.37	0.90	0.46	0.93
pasear	0.63	0.54	0.52	0.80	0.68	0.87
rama	0.06	0.08	0.11	0.20	0.05	0.00
humo	0.07	0.00	0.13	0.20	0.09	0.47
asomar	0.02	0.00	0.01	0.00	0.00	0.00
sucio	0.25	0.62	0.37	0.60	0.58	0.87
escalera	0.41	0.23	0.58	0.40	0.76	0.67
obrero	0.00	0.00	0.04	0.10	0.01	0.20
rueda	0.12	0.08	0.15	0.30	0.20	0.13
bolsillo	0.03	0.08	0.11	0.10	0.14	0.40
dulce	0.50	0.54	0.79	0.90	0.95	0.80
perdido	0.31	0.46	0.54	0.70	0.50	0.67
borrar	0.16	0.54	0.32	0.50	0.51	0.73
adivinar	0.03	0.38	0.06	0.70	0.27	0.67
cárcel	0.05	0.15	0.11	0.70	0.07	0.53
alcalde	0.00	0.15	0.03	0.50	0.02	0.20

rezar	0.04	0.23	0.04	0.10	0.01	0.13
ancho	0.05	0.00	0.14	0.40	0.05	0.07
orilla	0.15	0.08	0.09	0.10	0.17	0.00
pecado	0.12	0.00	0.29	0.10	0.20	0.00
alegre	0.57	0.92	0.65	0.80	0.85	0.87
sudor	0.05	0.00	0.05	0.30	0.02	0.27
huella	0.04	0.08	0.03	0.20	0.01	0.20
pisar	0.35	0.00	0.30	0.20	0.27	0.00
vientre	0.18	0.00	0.09	0.00	0.07	0.07
pálido	0.03	0.00	0.05	0.30	0.04	0.20
socio	0.37	0.08	0.45	0.60	0.50	0.60
gota	0.16	0.00	0.14	0.20	0.18	0.13
castigar	0.06	0.23	0.12	0.30	0.10	0.07
tonto	0.45	0.46	0.55	0.50	0.55	0.67
picar	0.13	0.00	0.32	0.30	0.10	0.13
orgulloso	0.10	0.23	0.23	0.50	0.30	0.47
suelto	0.09	0.00	0.20	0.00	0.13	0.00
eje	0.03	0.00	0.04	0.00	0.03	0.00
apellido	0.88	0.92	0.47	0.90	0.59	0.93
docente	0.03	0.15	0.06	0.20	0.04	0.13
colega	0.05	0.23	0.06	0.70	0.05	0.73
cerro	0.33	0.08	0.41	0.20	0.27	0.07
uña	0.49	0.31	0.50	0.50	0.53	0.13
chiste	0.14	0.15	0.38	0.50	0.34	0.60
desafío	0.02	0.46	0.05	0.30	0.07	0.27
penumbra	0.04	0.00	0.07	0.00	0.10	0.00
resaltar	0.09	0.00	0.17	0.20	0.16	0.07
ahogar	0.05	0.00	0.11	0.00	0.15	0.20
hueco	0.04	0.08	0.08	0.10	0.08	0.07
sabiduría	0.01	0.08	0.00	0.20	0.01	0.00
tibio	0.09	0.00	0.07	0.20	0.09	0.00
cumbre	0.03	0.00	0.06	0.00	0.02	0.00