

# HOW MANY COLLOCATIONS DO HERITAGE SPEAKERS KNOW? THE EFFECTS OF LINGUISTIC AND INDIVIDUAL VARIABLES

---

EVE ZYZIK

**ABSTRACT:** The current study examines a previously understudied dimension of heritage speakers' lexical knowledge by focusing on verbal collocations. Two tests were designed in order to assess both receptive (recognition) and productive (recall) knowledge of sixty Spanish collocations. The collocations were divided into three types (congruent, partially congruent, and incongruent) based on a ratings survey that established their degree of correspondence with English. Participants' language dominance and their use of Spanish in various daily activities were included as individual variables. The results indicate that the participants knew a vast majority of the collocations on the recognition test, but that their ability to recall the collocations was somewhat more limited. Congruency had a significant effect on participants' performance, but this finding must be interpreted in light of the interaction between congruency and word frequency. Significant correlations were found between performance on both tests and language dominance, as well as a number of variables involving interaction in Spanish (text messaging) and exposure (listening to music, reading for fun). These data are discussed in relation to previous studies on the acquisition of collocations and heritage speakers' knowledge of individual words.

**KEYWORDS:** *vocabulary, collocations, frequency, language dominance, individual differences*

---

EVE ZYZIK, University of California, Santa Cruz  
e-mail: ezyzik@scsc.edu

---

## 1. INTRODUCTION

A central piece of our lexical knowledge is knowing which words commonly go together to form sequences such as *human rights*, *densely populated*, and *take a break*. These sequences, deemed collocations, are generally defined as the statistical tendency of two or more words to co-occur in discourse (Schmitt, 2000). From a psychological point of view, collocating words predict each other and thus, form an association that eventually becomes consolidated in long-term memory as a single unit (Ellis, 2001). It is widely acknowledged that collocations are indispensable to producing language fluently and idiomatically. For this reason, they have been extensively researched in the context of second language (L2) learners, who often struggle with knowing which words pattern together despite having relatively advanced language proficiency. To date, many researchers (cf. Foster, 2009; Nesselhauf, 2003, Siyanova & Schmitt, 2008) have shown that L2 learners lack the range of collocations available to native speakers, both in terms of quantity (using fewer collocations overall) and quality (using them accurately and appropriately). In fact, a recent study by Dąbrowska (2019) found that a test of collocations yields the largest differences between native speakers and L2 learners, more so than tests of single word vocabulary and grammar.

In contrast to the abundant research with L2 learners, there is virtually no research that examines collocational knowledge among heritage speakers (see Treffers-Daller, Daller, Furman, & Rothman, 2015, for a notable exception). For Spanish in particular, there is some discussion of collocations in studies that examine calques (cf. Fairclough, 2013, Ortigosa, 2010), but for the most part, our understanding of Spanish heritage speakers' vocabulary knowledge is limited to individual words (see Fairclough & Garza, 2018, for a recent review). Unfortunately, in neglecting multi-word sequences such as collocations, we run the risk of seriously underestimating heritage speakers' vocabulary knowledge. Thus, the main aim of this study was to expand the research base in this area by examining the performance of Spanish heritage language learners (HLLs) on two tests of collocational knowledge that measured their ability to recognize and produce collocations. The stimuli include a representative sample of verbal collocations such as *recaudar fondos* ("to raise money") as well as collocations with an intervening preposition such as *tomar en serio* ("to take seriously"). The selection of the stimuli draws on corpora of Spanish,

frequency-based dictionaries, and surveys administered in the community to first-generation speakers. The design of the tasks relies on established methodologies from the field of second language acquisition (SLA).

## 2. BACKGROUND

### 2.1 Previous Research on Collocations

There is a vast literature on collocations in the field of SLA, with the bulk of this work being done on English collocations (see Boers & Webb, 2018, for an overview). Many studies involve a comparison between native speakers and L2 learners by extracting the collocations from a corpus of written samples (cf. Durrant & Schmitt, 2009; Laufer & Waldman, 2011). Intervention studies have compared different learning conditions to determine how collocations are learned from exposure (cf. Sonbul & Schmitt, 2013; Szudarski & Carter; Toomer & Elgort, 2019). Classroom-based research such as Boers, Demecheleer, Coxhead, and Webb (2014) and Peters (2016) have measured learning gains that take place after students work with different types of collocation exercises. Finally, psycholinguistic studies (cf. Sonbul, 2015; Wolter & Yamashita, 2015) have measured native and non-native speakers' processing of collocations in laboratory settings.

To my knowledge, there is only one published study that was specifically conducted to explore collocations among heritage speakers. Treffers-Daller et al. (2016) examined collocations among Turkish heritage speakers living in Germany and also those that had returned to Turkey (returnees). The researchers collected oral story retellings and analyzed the participants' use of collocations containing the verb *yap-* ("to do") as in *ders yap-* ("to do a lesson"). Their pioneering research shows that heritage speakers in Germany have created a new set of noun-verb collocations with *yap-*, which are different from the conventional patterns of Turkish speakers in Turkey, who tend to use the verb *et-* or other more specific verbs. What is more, the group of returnees sheds new light on the learning of collocations from exposure to monolingual input: after a certain period of time (one year), the returnees recognize that *yap-* is not the conventional choice in Turkey and after seven years they are no longer significantly different from monolingual users of Turkish.

Additional information about the use of collocations among heritage speakers can be gleaned from studies on calques, which are a manifestation of cross-linguistic influence. In the U.S. context, calques are a type of semantic extension that occurs when an English meaning is mapped onto an existing Spanish word. For example, the Spanish noun *carpeta* (“folder”) takes on the English meaning of “carpet” due to their phonological similarity. Calques can also affect multi-word units such as *jugar la guitarra*, which would compete with the Spanish collocation *tocar la guitarra*. Ortigosa (2010) presents an overview of these multi-word calques or *calcos fraseológicos* in a corpus of U.S. Spanish, arguing that they constitute linguistic innovations because they create new collocations in Spanish that did not exist before. Fairclough (2013) investigated calques with a translation exercise that specifically targeted words and phrases susceptible to English influence. The HLLs who completed the translation exercises generally expressed “to make a decision” with *hacer* rather than *tomar*. Thus, it seems that *tomar una decisión* is losing ground to the English-influenced *hacer una decisión*. Fairclough concludes that some of the items are already well established in the Spanish of the United States.

It is important to note that studies such as Fairclough (2013) were not designed to examine the full collocation repertoire of HLLs, but rather to understand the nature of English influence on the lexicon of these speakers. It is possible that HLLs utilize some English-influenced collocations (e.g., *hacer una decisión*) while at the same time having knowledge of hundreds of other Spanish collocations. In other words, examining the collocational knowledge of Spanish HLLs is a much broader question than looking specifically at cases of lexical transfer.

This review of the research would be incomplete without mention of classroom-based studies that aim to teach Spanish collocations to L2 learners. Stengers and Boers (2015) conducted a study with Dutch learners of Spanish who were at the A2 level of proficiency according to the Common European Framework of Languages (CEFR). They selected thirty-two verbal collocations such as *levantar la voz* and *llamar la atención* that were unknown on pilot testing with another cohort of students. Stengers and Boers compared two treatment conditions, both of which involved having learners complete fill-in-the-blank exercises with the verb missing. The treatment did not result in major learning gains; in both conditions the gains were less than 20% and the post-test scores were less than 50%. Another intervention study is that of

Jensen (2017), who compared three treatment conditions for teaching twenty-five Spanish collocations to a group of L2 English speakers. Jensen selected the verb-noun collocations from course readings, which allowed him to compare explicit instruction with an implicit approach (i.e., simply exposing learners to the collocations through reading). Both explicit groups improved significantly in their knowledge of collocations whereas the implicit group made almost no improvement. It is worth noting that Jensen used a multiple-choice test that presented intact collocations as options; in other words, learners did not have to combine words to make a collocation (like in Stengers and Boers' study). Clearly, the testing format has an impact on learners' performance and can complicate comparisons across studies. The next section discusses various options of measuring learners' knowledge of collocations.

## 2.2 Methods of Assessing Collocational Knowledge

One of the pressing issues for researchers who study collocations is how to measure learners' knowledge of these multi-word sequences, which has proven to be more difficult than testing individual words (Gyllstad & Schmitt, 2019). Given the variety of aims and approaches to studying collocations with L2 learners, this section will focus on studies that have a point of methodological comparison to the current one. With respect to receptive knowledge, Nguyen and Webb (2016) developed a comprehensive test of English collocations based on word frequencies. They focused on verb-noun and adjective-noun collocations in a multiple-choice format. For example, given the noun "money" participants had to choose between four verbs (check, drop, make, or miss) that form a natural collocation. In this testing format, the node word is "money" and the collocate is "make". The participants in this study were Vietnamese learners of English who had studied at least seven years of English as a foreign language (EFL). Their results indicate that these L2 learners have weak knowledge of collocations overall (less than 50% accuracy), but that performance was higher on the most frequent noun-verb and adjective-noun collocations. It is important to note that the test developed by Nguyen and Webb is one that targets form recognition, that is, participants' ability to recognize a given combination of words as a natural phrase. It does not measure participants' comprehension of collocations or their ability to use them in an appropriate way (see Macis and Schmitt, 2017, for a test that involves meaning recall).

In addition to form recognition, researchers have also tested participants' ability to produce a target collocation in response to a stimulus. An exemplary study in this regard is González Fernández and Schmitt (2015), who designed a test of productive knowledge targeting 50 English collocations. Since their participants were native Spanish speakers (EFL learners), the test was given in a bilingual format with a context in Spanish followed by the main sentence in English. The main sentence contains blanks for the target collocation, but the first letter of each word is provided as a way of constraining the range of responses. For example, for the target collocation "foreign accent", the main sentence looks like this: "Yes, he spoke with a bit of a f\_\_\_\_\_ a\_\_\_\_\_." (González Fernández & Schmitt, p. 125). This type of test is described as requiring the most advanced degree of knowledge (Laufer & Goldstein, 2004).

In sum, the studies reviewed thus far have tested different dimensions of collocational knowledge. Form recognition can be tested in a multiple-choice format, with participants having to match a node word with its collocater (cf. Nguyen & Webb, 2016). The ability to produce collocations (form recall) has been tested using a c-test format in which the initial letters of target words are provided (cf. González Fernández & Schmitt, 2015). This last test avoids guessing effects, but it does not necessarily imply that participants can use the collocations spontaneously in their speech or writing.

### 2.3 Variables in Collocation Research

In addition to the testing format, researchers generally consider a number of linguistic variables when selecting stimuli such as the frequency of the collocation and the mutual information (MI) score. MI is a measure of collocation strength or exclusivity: larger MI values indicate that the two words appear predominantly in each other's company. Gablasova, Brezina and McEnery (2017) explain that higher MI scores tend to favor lower frequency collocations as well as specialized or technical terms. For example, "zig zag" may have a very high MI score but a low overall frequency in a given corpus. In addition to MI, researchers have controlled for the frequency of the component words, semantic transparency, and congruency. A given study generally focuses on one or two of these variables rather than trying to account for all of them in a single design. For reasons of space, only two

factors will be discussed in more detail here: congruency and the frequency of the component words.

It is crucial to establish the frequency of the component words because lack of knowledge of a particular collocation may be due to the fact that the learner does not have one (or both) of the items in his/her lexicon (see Gablasova et al., 2017). Nguyen and Webb (2017) indicate that node word frequency was the most important predictor of collocational knowledge, explaining 12.1% of the variance. Similarly, Stengers and Boers (2015) tested verb frequency as a predictor of learning Spanish verb-noun collocations. They distinguished between verbs that were very high-frequency (among the most frequent 1000 words in Spanish) and those that were not. Their results confirm that verb frequency was a significant predictor of learning. In the current study, frequency of both component words (nouns and verbs) will be considered as an independent variable.

Another variable that has been considered is congruency between languages, which is defined as the presence or absence of a literal translation equivalent (Nesselhauf, 2003). For example, “to make sense” is expressed as *tener sentido* (to “have” sense) in Spanish. Although they are synonymous, these collocations use a different verb in each language and thus would be considered incongruent. In contrast, the collocation “to live in peace” is expressed identically in Spanish (*vivir en paz*) and thus would be classified as congruent. In other words, congruent collocations can be translated from one language to the other using a word-for-word strategy (cf. Pulido & Dussias, 2020). Although this seems like a clear criterion, a binary classification of congruency is often quite difficult in practice. Peters (2016) explains that “congruency might be less easy to operationalize than previously hypothesized because of polysemy and prototypicality of meaning” (pg. 130). Some of these problems can be mitigated with a three-way classification of congruency (e.g., incongruent, partially congruent, and congruent), which was done by Nguyen and Webb (2016). Another methodological concern is the use of a single rater (often the researcher) to determine whether a given collocation is congruent or not. A different approach would be to administer a ratings survey to a larger number of bilingual speakers in order to measure the degree of congruency between Spanish and English collocations. A similar methodology has been used in previous studies for the purpose of establishing semantic similarity between morphologically related words

(Gonnerman, Seidenberg, & Andersen, 2007). A ratings survey provides valuable information above and beyond the researcher's individual assessment of the linguistic variable at hand.

In addition to the abovementioned linguistic variables, researchers have studied a host of individual variables related to learners' engagement with the target language. As explained by González Fernández and Schmitt (2015), the acquisition of collocations goes beyond mere exposure to language and depends on "the kind of high-quality engagement with language that presumably occurs in a socially integrated environment" (pg. 98). In the same study, González Fernández and Schmitt found significant correlations between L2 learners' knowledge of collocations and their exposure to the target language in the form of reading, watching TV/movies, and using social media. In a study with Chilean university students, Macis and Schmitt (2017) found significant effects for time spent studying in an English-speaking country and time spent reading on participants' knowledge of target collocations with figurative meanings such as "drop the ball". If we assume that the prototypical HLL has many opportunities for meaningful interaction with Spanish speakers, these language use variables take on additional relevance. In the current study, HLLs' personal engagement with Spanish (e.g., text messaging, social media, reading, music, etc.) will be assessed via self-report.

### **3. THE CURRENT STUDY**

#### **3.1 Research Questions**

The design of the study was guided by three primary research questions:

1. To what degree do Spanish HLLs recognize and produce a diverse set of frequent collocations?
2. Which linguistic variables (e.g., congruency with English, frequency of component words) predict which collocations are known among Spanish HLLs?
3. How do individual factors (e.g., language dominance, personal engagement with Spanish) relate to participants' knowledge of collocations?



### 3.2 Item Development

The process of selecting collocations for inclusion in an experimental study is always a challenge, as it is difficult to compile a sample that truly represents the vast number of collocations in Spanish (or any language). Nevertheless, by following precise methodological steps, it is possible to develop a list of collocations that responds to certain criteria. The sampling process in the current study followed four steps:

1. Initial sampling from a frequency-based dictionary of collocations (Robles-Saez, 2011).
2. A rating survey to determine how congruent the Spanish collocations are with English.
3. A community survey to limit the list to collocations known by first-generation speakers.
4. Confirmation of frequency and MI score (based on the Corpus del Español).

The sampling process began by consulting Robles-Sáez (2011), who compiled a dictionary of 3000 collocations informed by data from the Corpus de Referencia del Español Actual (CREA) and also by another, smaller corpus of oral conversations among Spanish speakers in the United States.<sup>1</sup> The collocations are all verbal (i.e., a verb followed by a noun, an adverb, or a prepositional phrase). From the 3000 collocations in Robles-Sáez's dictionary, I selected collocations that varied in terms of the main verb, focusing on collocations that were described as "neutral" or "informal" in the dictionary rather than those that were more literary (e.g., *turbar el sueño*) or highly technical (e.g. *implantar silicona*). This initial selection process resulted in an initial list of 103 collocations.

The second step in the sampling process was a ratings survey in order to determine the degree of congruency between the 103 collocations and their English equivalents. The ratings survey was administered via Qualtrics and distributed via anonymous links to emails and social media. The participants were all Spanish-English bilinguals (with no requirement for being a native speaker of one or the other) and were mostly faculty and graduate students at

various universities. A total of 72 individuals completed the ratings survey; half identified their native language as English ( $n = 36$ ), slightly less than half ( $n = 30$ ) identified their native language as Spanish, and six participants marked “both Spanish and English.” Instructions were provided to define collocations and to explain congruency as a “match” between English and Spanish expressions (with examples). Four practice items were included to familiarize participants with the rating scale, which ranged from 1 (incongruent) to 5 (congruent). Participants were encouraged to use the entire scale for items that were somewhere in between.

The results of the ratings survey confirmed that some collocations are perceived as highly congruent (with mean scores above 4), some are incongruent (mean scores below 2) and many others fall somewhere in between. Table 1 shows a sample of the results, with congruency mean scores in parentheses. The next phase in item selection was a community survey administered to first-generation Spanish speakers in the state of California, representing both urban and rural areas. The goal of the community survey was to eliminate any collocations that are not used by these speakers, and thus, probably not part of the input of heritage speakers. A total of 18 participants completed the community survey; their average age was 44.14 years and the average time spent in the U.S. was 22.11 years. These surveys were administered orally in order to minimize the effects of literacy and educational attainment. The instructions (also read orally) explained that the goal of the survey was to find out which expressions were common in the community; participants were instructed to respond “yes” to the expressions that they used and “no” to the ones they didn’t use. If they didn’t like a particular expression, they could give alternatives or comment further on what they might say instead (this was not required). Their responses were recorded on a printed checklist.

**Table 1.** Congruency Based on Ratings Survey

Congruent	Partially congruent	Incongruent
aceptar las disculpas (4.28)	atar cabos (2.97)	levantar cabeza (1.33)
cometer un delito (4.6)	cambiar de opinión (2.92)	pasar hambre (1.58)
leer entre líneas (4.69)	cumplir una promesa (3.13)	pasar de largo (1.74)

**Note:** Congruency was rated on a 1–5 scale. Mean ratings are given in parentheses.

The responses from the community surveys were analyzed with the goal of eliminating any problematic or highly variable collocations. Collocations that were not used by five or more speakers were eliminated. Other collocations were eliminated because participants proposed alternatives, and it could be that these alternatives are more frequent in the community. For example, for *seguir el hilo*, eight participants proposed *seguir la corriente* instead. The end of this process was a final list of 60 collocations that are representative of the usage of first-generation speakers in various communities in the state of California. These collocations were grouped into three categories based on the results of the ratings survey, as shown in Table 1.

These 60 collocations were checked for frequency by consulting the Corpus del Español (Web/Dialects). All collocations had at least 300 occurrences in the corpus (lemmatized frequency). Note that lemmatized frequency reflects all possible conjugations of the verb (e.g., *tocar fondo*, *toque fondo*, *tocó fondo*, *tocamos fondo*, etc.). The MI scores of the collocations ranged from 3.1 to 12.84 (median MI score = 6.41). This is consistent with the recommendation that refers to a minimum MI score of 3.0 for defining collocations (Hunston, 2002).

Finally, the component words in each collocation were checked for frequency using the Word and Phase section of the Corpus del Español (<https://www.wordandphrase.info/span/>). This tool allows lemma searches and provides detailed information including raw frequency, rank frequency (up to 40,000) and distribution by genre (spoken, fiction, news, and academic). Rank frequency is especially useful since we want to know if any of the component words fall outside of the high-frequency range. In the current study, a conservative approach was taken to identify those collocations that had at least one component word beyond the 2000 frequency level. Twenty-one collocations were in this category, having either a lower-frequency verb (e.g., *colmar* in *colmar la paciencia*) and/or a lower-frequency noun (e.g., *riendas* in *tomar lasriendas*). The remaining thirty-nine collocations consist of high frequency component words within the top 2000 lemmas. Appendix A lists all of the collocations along with the corresponding frequency information.

### 3.3 Methodology

#### 3.3.1 *Participants*

The participants were thirty heritage speakers of Spanish who were enrolled at a mid-size public university in California. There were 20 females, 9 males, and one non-binary participant. The average age of the participants was 20.43 with an age range of 18 to 27 years. The majority of the participants ( $n = 23$ ) were born in the United States whereas seven of them were born either in Mexico or Central America. For the foreign-born participants, the age of arrival in the U.S. ranged from 0 to 7 years of age (average age of arrival = 3.28 years).

With respect to their parents' background, the large majority of the participants ( $n = 27$ ) indicated that both of their parents were born in Mexico or Central America (El Salvador or Guatemala). There were only three participants who indicated having one U.S.-born parent. Among the parents, 53% had completed at least high school. A smaller percentage (18%) had completed college or graduate school.

#### 3.3.2 *Questionnaires*

Participants completed a biographical questionnaire based on Gatti and O'Neill (2017), which elicited information about their family background, their prior language learning experience, and their current use of Spanish and English for various purposes. Nine questions elicited information about participants' current use of both languages for text messaging, posting on social media, watching TV, listening to music, listening to talk radio (e.g., news, podcasts), reading books, reading articles (including newspapers, magazines and blogs), and reading for fun (e.g. Facebook, memes, message boards). For each question, participants indicated the frequency with which they engage in this activity using Spanish and using English. The response options were: never, rarely, sometimes, usually, always, or not applicable. The data were coded with numerical values (0 = never, 1 = rarely, 2 = sometimes, etc.), which constitute an ordinal scale. These language use variables will be

analyzed in order to understand how they relate to participants' knowledge of collocations.

In addition to the biographical questionnaire, participants completed the Bilingual Language Profile (Birdsong, Gertken, & Amengual, 2012), which yields a score of language dominance. Positive scores indicate English dominance, negative scores indicate Spanish dominance, and scores near zero indicate more balanced bilingualism. Rather than grouping participants into discrete groups, the Bilingual Language Profile (BLP) scores will be used in subsequent analyses as a continuous variable. Ortega (2020) and Leal (2018) caution against transforming variables that are inherently continuous into categorical ones. Since language dominance is conceptualized as gradient (e.g., English dominance involves various degrees of being dominant in that language), I have opted to retain it as a continuous variable.

### *3.3.3 Collocation tests*

Two tests were utilized in the current study. The first was a form recognition test based on Nguyen and Webb (2016), but with some modifications. The form recognition test employs a multiple choice format in which the verb is provided and participants choose among three options that co-occur most naturally with it. After selecting a response (a, b, or c), participants were asked to indicate their level of confidence in their response using a three-way scale.<sup>2</sup> This test was administered on a large screen via PowerPoint slides with time pressure: each experimental item remained on the screen for 10 seconds. Participants recorded their responses on a printed answer sheet. The instructions, which aimed to avoid metalanguage (e.g., verb, noun) as much as possible, were provided in English:

You will see a prompt word in ALL CAPS on the left side of the screen. Choose the word that combines most naturally with the prompt word. There will be three options; you should choose only one. Note: The word you choose should combine directly with the prompt word.

The instructions were followed by three practice items (untimed) and then the sixty experimental items (timed). The format of the test is shown below.

- ROMPER    a. las reglas  
               b. las dificultades  
               c. las desgracias

As noted by Nguyen and Webb (2016), a key consideration in the design of such a test are the distractors because these have to be unnatural combinations of words. Accordingly, a distractor bank was compiled consisting of masculine and feminine nouns, adjectives, and adverbs from the top 5000 words in Spanish (according to the Corpus del Español, Word and Phrase search). From the distractor bank, the following criteria were used in selecting distractors for each item:

- 1) The distractors had to be the same part of speech, gender, and number as the correct answer.
- 2) When the distractors and the node word were combined, their frequency in the Corpus del Español was less than 5.

The form recognition test was piloted with two native speakers of Spanish in order to flag any poor distractors. Based on pilot testing, six items were revised by substituting distractors that were potentially problematic.

The second instrument was a form recall test based on the methodology of González Fernández and Schmitt (2015). Participants are asked to supply the missing part of a collocation after reading a brief context that cues meaning. In the current study, the contexts were provided in English, followed by a Spanish sentence with an incomplete collocation, as shown in example (1):

- (1) After the corruption scandal, the president's popularity is at an all-time low.  
       Está claro que el presidente está en su peor momento; su popularidad ha *tocado* f\_\_\_\_\_ . [complete collocation: *tocado fondo*]

To constrain the possible responses on this task, the first letter of the missing word is provided. The task designed by González Fernández and Schmitt was more difficult in the sense that both parts of the collocation were missing

(i.e., participants had to supply both the verb and the noun). In the current study, however, the decision was made to provide the verb and any prepositions (if relevant), leaving only one blank to fill. This decision was based on pilot testing with native speakers, who found the task to be very difficult if multiple elements of the collocation were missing. Finally, a choice of determiners were provided for those collocations with determiner plus noun, as in example (2):

- (2) It's best to resolve situations in which a person owes money to a family member.

Francisco vendió una de sus casas para *saldar el/la d \_\_\_\_\_* que tenía con su hermano. [complete collocation: *saldar la deuda*]

The design of the form recall test underwent several rounds of pilot testing with six native speakers from various countries (Spain, Costa Rica, and Mexico). The pilot testing was critical in making sure that the contexts were clear, but did not provide the target collocation in English (this is especially important in the case of congruent collocations). Based on pilot testing, changes were made to several items in order to make them sound more natural. The research instruments are summarized in Table 2. Note that the same sixty collocations were used in both tasks, as described in the earlier section (“Item development”).

**Table 2.** Research Instruments

Form recognition (receptive knowledge)	Form recall (productive knowledge)
60-item multiple choice test with one correct answer and two distractors. Verb is provided; participants choose the word that most naturally combines with the verb. Time pressure.	60-item fill-in-the blank test with context in English and target sentence in Spanish. Collocation is incomplete; verb is provided as well as the first letter of the missing item. No time pressure.

Let us consider the types of knowledge that are measured by both tests (see González Fernández and Schmitt 2019 for more information about different measures of word knowledge and how they relate to each other). The form recognition test measures participants’ ability to recognize a given collocation from a number of options. On the other hand, the form recall test measures the ability to retrieve the target collocation from memory based on

a stimulus (in this case, context clues). When used in combination in a single study, these tests can give us a more comprehensive picture of participants' collocational knowledge. Crucially, the form recognition should be administered after the form recall test in order to minimize potential interference from one test to the other.

### 3.3.4 Procedure

The research was carried out in small group sessions with 6 to 12 participants in each session. After being welcomed and signing the informed consent form, participants filled out the biographical questionnaire. Then they began the form recall test. Afterwards, they completed the BLP. When all participants had completed these steps, the form recognition test was administered by projecting the slides on a large screen. This was the only timed portion of the study (recall that each slide remained on the screen for ten seconds). All the remaining activities were untimed, allowing participants to work at their own pace. The total time for completing all the tasks ranged from 45 to 65 minutes. Participants received \$20 in compensation upon completion of the research session.

### 3.3.5 Data coding and analyses

Both tests were scored dichotomously (1 point awarded for a correct response; 0 points for an incorrect response). On the form recall test, spelling errors were common on some of the items (e.g., *hicieron las pazes* (paces), *sufrir las consecuencias* [consecuencias]). All responses with spelling errors that did not drastically alter the pronunciation of the target word were considered correct. Likewise, all written accent errors were ignored.

To answer the research question about the linguistic variables, a binary logistic regression was deemed appropriate since the dependent variable is dichotomous (e.g. a correct or incorrect response). The advantage of the binary logistic regression is that it is based on the number of observations and not on mean test scores. The current data set consists of 1800 observations on the form recall test and another 1800 observations on the form recognition test. In interpreting the results of the binary logistic regression, we consider the odds ratio ( $\exp(B)$ ), which predicts the odds in favor of an accurate response with a specific parameter of the independent variable.



To answer the research question about the individual variables, a series of correlations were computed with participants' mean scores. BLP scores (language dominance) were analyzed as a continuous variable. The individual variables related to personal engagement with Spanish (e.g., text messaging, reading for fun) were analyzed as ordinal variables.

## 4. RESULTS

### 4.1 Descriptive Statistics

The mean score on the form recall test for the participants as a group was 67.55%. The mean score on the form recognition test was higher: 88.17%. Table 3 reports the means, standard deviations, and ranges on both tests; the boxplots in Figure 1 show these data visually. Cronbach's Alpha was calculated as a measure of internal consistency for both tests. The form recall test yielded a Cronbach's Alpha reliability coefficient of .907; the reliability coefficient for the form recognition task was lower (.79) but still considered acceptable (George & Mallery, 2003).

There was a strong positive correlation between the mean scores on the recognition test and the form recall test:  $r(28) = .868$ , Pearson:  $p < .001$ .

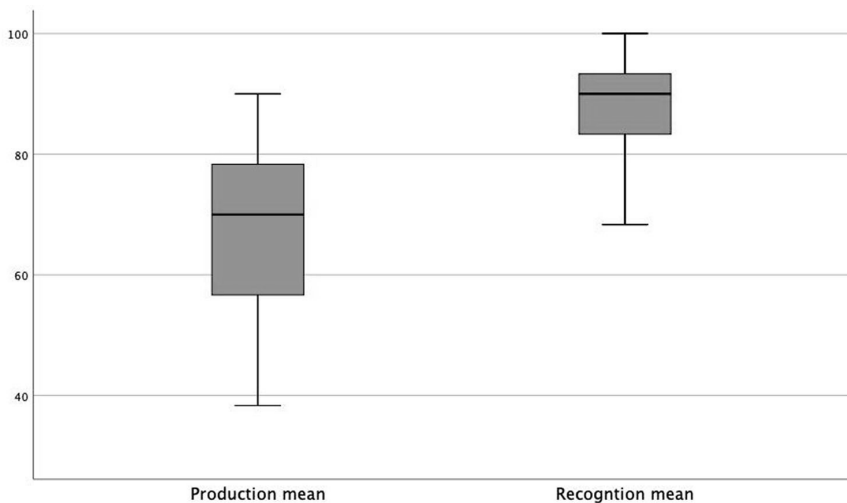


Figure 1. Mean scores on both collocation tests

**Table 3.** Descriptive Statistics

Form recall test (n = 30)	M = 67.55 (15.19)	Range: 38.33–90
Form recognition test (n = 30)	M = 88.17 (8.13)	Range: 68.33–100

## 4.2 Linguistic Variables

Beyond the group means reported above, we are interested in knowing whether congruency and word frequency affects the odds that a participant will recall or recognize a given collocation. Before presenting the statistical analysis, it is useful to see the distribution of correct responses with respect to item congruency. Table 4 shows the number of correct responses on both tests. The percentages are calculated based on the total number of items in each congruency category. The data in Table 4 reveals that incongruent items result in lower accuracy on both tests. Reading across the bottom row of the table, we see that congruent items on the recognition test are almost at ceiling (94%). In contrast, congruent items on the recall test have a lower accuracy rate (71%), and in fact, the “partially congruent” items yielded a higher number of correct responses. This suggests that congruency does not have a straightforward effect on participants’ performance on the form recall test.

A binary logistic regression based on 1800 observations (30 participants multiplied by 60 target items) was conducted separately for each test. Congruency was coded as a categorical variable with three levels: incongruent, partially congruent, and congruent. Frequency of the component words was included with two levels: low and high. In this model, ‘incongruent’ was selected as the baseline, meaning that the other levels of congruency were compared to it. Likewise, ‘low frequency’ was the baseline for comparison. The results of the binary logistic regression are shown in Table 5 (recall test) and Table 6 (recognition test).

**Table 4.** Cross Tabulation of Responses by Congruency

	Correct responses: Recall	Correct responses: Recognition
Incongruent (n = 600)	324 (54%)	478 (80%)
Partially congruent (n = 600)	468 (78%)	545 (91%)
Congruent (n = 600)	424 (71%)	564 (94%)

For the form recall test, the logistic regression model was statistically significant,  $\chi^2(5, N = 1800) = 169.334, p < .001$ . With the two predictors, the model explained 12.5% of the variance (Nagelkerke  $r^2$ ). The interaction between congruency and word frequency is significant ( $p < .001$ ). When congruency is partial and word frequency is high, the odds ratio,  $\text{Exp}(B) = 3.743$ . A similar pattern is revealed for the congruent collocations when these contain high frequency component words: the odds ratio,  $\text{Exp}(B) = 3.526$ . This means that participants are 3.5 times more likely to produce an accurate response when both conditions are met (i.e., the collocation is congruent and the component words are of high frequency).

We turn now to the form recognition test. The logistic regression model was statistically significant,  $\chi^2(5, N = 1800) = 82.82, p < .001$ . With the two predictors, the model explained 8.7% of the variance (Nagelkerke  $r^2$ ). Congruency predicted the odds of a correct response: when congruency was partial, the odds ratio,  $\text{Exp}(B) = 1.71, p < .001$ . For the congruent items, the odds ratio,  $\text{Exp}(B)$ , increases to 2.67,  $p < .001$ . Word frequency was not a significant predictor in this model. Since the interaction between word frequency and congruency was not significant ( $p = .075$ ), we can interpret the main effect for congruency.

**Table 5.** Binary Logistic Regression (Form Recall Data)

Step 1a	B	S.E.	Wald	df	Sig.	Exp(B)
Congruency			7.476	2	.024	
Congruency (1)	.386	.165	5.480	1	.019	1.471
Congruency (2)	-.138	.170	.659	1	.417	.871
Frequency (1)	.045	.167	.072	1	.789	1.046
Congruency*Frequency			85.561	2	.000	
Congruency(1)*Frequency(1)	1.260	.204	38.112	1	.000	3.526
Congruency(2)*Frequency(1)	1.320	.192	47.449	1	.000	3.743
Constant	.160	.082	3.832	1	.050	1.74

a. Variable(s) entered on step 1: Congruency \* Frequency

**Note:** Congruency (1) refers to partially congruent; Congruency (2) refers to congruent. Frequency (1) refers to high frequency component words.

**Table 6.** Binary Logistic Regression (Form Recognition Data)

Step 1a	B	S.E.	Wald	df	Sig.	Exp(B)
Congruency			12.664	2	.002	
Congruency (1)	.540	.248	4.753	1	.029	1.717
Congruency (2)	.984	.292	11.354	1	.001	2.676
Frequency (1)	.262	.205	1.643	1	.200	1.300
Congruency*Frequency			5.168	2	.075	
Congruency(1)*Frequency(1)	.709	.352	4.045	1	.044	2.031
Congruency(2)*Frequency(1)	.646	.402	2.584	1	.108	1.909
Constant	1.213	.154	62.382	1	.000	3.364

a. Variable(s) entered on step 1: Congruency \* Frequency

**Note:** Congruency (1) refers to partially congruent; Congruency (2) refers to congruent. Frequency (1) refers to high frequency component words.

### 4.3 Individual variables

#### 4.3.1 Language dominance

The BLP scores for the participants in the current study ranged from  $-78.55$  to  $89.09$ . To determine the relationship between collocational knowledge and language dominance (both continuous variables), two Pearson correlations were calculated. These data are summarized in Table 7. The data reveal a strong negative correlation between scores on the form recall test and language dominance. In other words, the participants who were more English-dominant scored lower on this test. The findings for the form recognition test show a similar relationship, although the correlation is considered to be moderate.

**Table 7.** Correlations between BLP Scores and Performance on Both Tests

	BLP Scores (language dominance)
Form recall test	$r = -.719, p < .001$
Form recognition test	$r = -.564, p < .001$

### 4.3.2 Personal engagement with Spanish

The data from participants' use of Spanish for nine different activities were analyzed: text messaging, posting on social media, writing emails, watching TV, listening to music, listening to talk radio, reading books, reading articles, and reading for fun. To what extent do these language use variables relate to participants' knowledge of collocations? The correlations between each variable and the mean scores on the form recall and form recognition tests are shown in Table 8.

From Table 8, we note that four of the language use activities have a positive relationship with collocational knowledge: text messaging, posting on social media, listening to music, and reading for fun. Subsequently, these four variables were combined to create a composite variable called "personal engagement with Spanish" (see González Fernández & Schmitt, 2015). The correlation between the composite variable and scores on the recognition test is positive,  $r(28) = .427, p = .003$ . Similarly, the correlation between the composite variable and scores on the production test is  $r(28) = .509, p < .001$ . This indicates that personal engagement with Spanish explains 18% of the variance on the form recognition test and 26% of the variance on the form recall test.

**Table 8.** Correlations between Individual Variables and Performance on Both Tests

Activity	Form Recall	Form Recognition
Text messaging	$r = .486^{**}$	$r = .442^{**}$
Posting on social media	$r = .320^*$	$r = .307^*$
Writing emails	$r = -.027$	$r = -.064$
Watching TV	$r = .289$	$r = .186$
Listening to music	$r = .333^*$	$r = .321^*$
Listening to talk radio	$r = .264$	$r = .193$
Reading books	$r = .132$	$r = .114$
Reading articles	$r = .199$	$r = .114$
Reading for fun	$r = .442^{**}$	$r = .329^*$

\*\*Kendall's tau  $p < .01$

\*Kendall's tau  $p < .05$

## 5. DISCUSSION

Based on the results of two tests (form recall and form production), we can make some preliminary observations about HLLs' knowledge of Spanish collocations. To answer the first research question, the mean scores show that HLLs scored near ceiling on a recognition test ( $M = 88.17$ ) of collocations and that they had somewhat weaker productive knowledge of the same items ( $M = 67.55$ ). This is consistent with the findings of recent vocabulary studies such as González Fernández and Schmitt (2019), who concluded that recognition knowledge is mastered before any type of recall knowledge. Moreover, they found that the different aspects of vocabulary were highly interrelated; in their data, recognition of collocations was significantly related to recall of collocations ( $r = .806$ ). The current study also found a strong positive correlation between scores on the two tests of collocational knowledge ( $r = .868$ ).

Returning to the mean scores of the HLLs in this study, readers might ask whether this should be considered a “good” level of knowledge of collocations. How we interpret these numbers depends inevitably on what kind of comparisons we make. In a recent overview of vocabulary knowledge among Spanish HLLs, Zyzik (2021) explains that different studies present seemingly contradictory findings: often HLLs are described as having robust vocabulary knowledge whereas other studies highlight their limited vocabulary. These apparent contradictions stem from several factors, including the kind of task that is used to assess vocabulary knowledge (i.e., receptive versus productive tasks) and who the comparison group is (i.e., L2 learners or monolingual native speakers).

Although direct comparisons are impossible, it is useful to situate the data from the current study in relation to previous research conducted with L2 learners. Macis and Schmitt (2017), who tested the knowledge of collocations among Chilean university students enrolled in graduate-level courses in English, found that knowledge of the meaning of the target collocations was found to be limited (33% accuracy). Boers and Stengers (2015) conducted an intervention study on Spanish verb-noun collocations, some of which were also tested in the current study (e.g. *levantar la voz* and *llamar la atención*). Their results indicate less than 50% accuracy on the post-test, despite the pedagogical treatment. González Fernández and Schmitt (2015) report a mean

accuracy of 56.6% on a form recall test among Spanish-speakers who had been studying English for an average of 13 years. Finally, Nguyen and Webb (2017) reported accuracy rates of 45% on a recognition test among Vietnamese EFL learners who had seven years of formal instruction. Taken together, these studies indicate that knowledge of collocations among L2 learners is often limited despite many years of study. Thus, viewed in the larger context of previous research on collocations, the data reported here suggest that HLLs have much stronger collocational knowledge than L2 learners.

In response to the second research question (linguistic variables), the results show that the effects of congruency are not the same across the two tasks. Congruency had an effect on the recognition test, with the odds in favor of congruent collocations resulting in a correct response. On the form recall task, however, the effect of congruency is conditioned by word frequency. The odds of providing a correct answer increased when the target collocation was congruent (or partially congruent) and when the component words were of high frequency. Word frequency did not have a significant effect on the recognition of collocations.

These results make sense if we consider the conservative way of defining frequency in this study: high frequency words were those within the 2000 most-frequent words in Spanish. The low frequency words (as defined in this study) were ranked above the 2000 threshold, but most of these were still within the 5000-frequency range. Fairclough (2011) has shown that HLLs tend to have receptive knowledge of the top 5000 words in Spanish. The results of the current study provide further support for this finding. There was not a significant effect for word frequency on the recognition test because it is likely that the participants had receptive knowledge of all (or almost all) the component words in the target collocations. Let us consider the mean scores of the seven collocations containing words beyond the 5000-frequency level (*surtir efecto*, *saldar*, *estirar*, *colmar*, *raya*, *riendas*, and *ridículo*). Of these, only *surtir efecto* was unfamiliar to the majority of the participants (47% mean accuracy on the recognition task). The remaining six collocations were known by 77% or more of the HLLs. In fact, *estirar las piernas*, *hacer el ridículo*, and *pasarse de la raya* were recognized correctly by more than 90% of the participants.

Overall, the current study found a facilitative effect for congruency, which is consistent with the results of previous research with L2 learners

(cf. Peters, 2016; Pulido & Dussias, 2020; Yamashita & Jiang, 2010; Wolter & Gyllstad, 2011). Nevertheless, the linguistic variables only explained 12.5% of the variance on the form recall test and even less of the variance (8.7%) on the recognition test. In other words, there are other factors that contribute to participants' knowledge of a given collocation. Not all incongruent collocations were equally difficult for the HLLs in this study. On the form recall test, some incongruent collocations yielded very low mean scores while others were well known: compare, for example, *partir de cero* (7%) and *pasar hambre* (90%). Conversely, some congruent collocations yielded low mean scores: *tomar las riendas* (37%) and *inspirar confianza* (23%). Congruency does not guarantee knowledge because the corresponding collocation in English might be unfamiliar to the participant (e.g., "take the reins"). Another explanation is that congruency is inherently subjective and that certain differences between collocations in two languages might not matter to highly proficient bilinguals. For example, *hacer el ridículo* is incongruent if we consider the conventional way of expressing this idea in English ("make a fool of yourself") or it could be perceived as partially congruent if it conjures up "be ridiculous" in the mind of the speaker. Peters (2016) reminds us that the congruency of a given collocation "might differ from learner to learner and will undoubtedly be affected by learners' proficiency and vocabulary size" (p. 130).

With respect to the third research question (individual variables), the data show that language dominance was significantly related to participants' knowledge of collocations. Participants who were more English dominant scored lower on the test of form recall ( $r = -.719$ ) and lower on the test of recognition ( $r = -.564$ ). The strong negative correlation between language dominance and the ability to recall collocations suggests that English dominance has a particularly detrimental effect on vocabulary production in Spanish. The BLP provides a composite score of relative dominance that takes into account language history, language use, self-perceived proficiency, and language attitudes. As such, it cannot be used to determine if certain types of interactions or activities contribute to participants' scores on a dependent variable. For this reason, another questionnaire was included to gauge HLLs' degree of engagement with Spanish, with targeted questions about daily activities such as text messaging, social media, listening to music, and reading for fun. These four variables combined (personal engagement with Spanish) explain 18% of the variance on the form recognition test and 26% of



the variance on the form recall test. In particular, text messaging in Spanish was significantly correlated to performance on both tests whereas more academically oriented activities, such as reading books and articles were not associated with increased knowledge of collocations. This interesting finding deserves further scrutiny to determine the nature of the text messages; it could be that text messaging in Spanish is a proxy for interacting with family and friends who are dominant in Spanish. If this is the case, it is not the mode of communication (text messaging) that contributes to HLLs' knowledge of collocations but rather the interaction with Spanish speakers.

## 6. LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

The findings reported here should be interpreted in light of the limitations of the study and the methodological choices in designing the research materials. First, the findings represent HLLs' ability to choose the noun that results in a felicitous combination with a given verb. Note that in both research tests, the verb was provided rather than the other way around. Research with L2 learners (cf. Stengers and Boers, 2015) has emphasized that the difficulty in verb-known collocations is selecting the correct verb because the verbs tend to be semantically light (e.g., do, make, have). In the current study, the decision was made to provide the verb in order to restrict the range of responses. Many collocations have more than one variant in which the verb differs (e.g., *meter / marcar un gol; aprobar / pasar un examen; poner / prestar atención*). In addition, it is the verb that is likely to exhibit cross-linguistic influence as in the case of English-influenced *hacer una decisión* (Fairclough, 2013). Future studies could investigate HLLs ability to supply the verb in verb-noun collocations, but care must be taken to keep the format consistent across tasks because of potential interference between recognition and recall tests.

A second limitation of the current study is that it did not control for the frequency of the collocations as a whole. Studies that have specifically investigated frequency as an independent variable have yielded mixed results. Durrant (2014) conducted a meta-analysis on this issue and found that frequency correlated only moderately with knowledge of collocations. In the current study, all of the collocations were relatively frequent according to their lemmatized frequency in the Corpus del Español. Moreover, relative frequency is dependent on register and modality (i.e., some collocations may

be less frequent in newspapers and academic discourse, but may be very common in informal conversations). For this reason, Gablasova et al. (2017) caution against using general corpora when trying to establish the connection between presumed input (exposure) and learners' collocational knowledge. Future studies may wish to explore this issue by controlling for the frequency of the target collocations within a particular subcorpus. That approach still leaves unanswered the lingering question of whether frequency information derived from a corpus reflects the actual exposure of HLLs. The methodology pursued in this study was to survey first-generation speakers from the community to ensure that all of the collocations were psychologically real combinations of words. This method could be expanded and improved upon by conducting a greater number of community surveys and asking more detailed questions about frequency of use.

Finally, the robust association between participants' productive knowledge of collocations and language dominance (BLP) suggests that the form recall test could be useful as a diagnostic tool in HLL research. Researchers have debated the validity of the DELE as an assessment for HLLs (Carreira & Potowski, 2011; Van Osch et al, 2018), but there are currently few alternatives. The form recall test developed for this study had a wide range of scores (33–90) and a high degree of internal consistency ( $\alpha = .91$ ) This measure would likely correlate with other aspects of linguistic proficiency, including vocabulary size (see Dąbrowska, 2019). More empirical research is needed to determine the relationship between knowledge of collocations and other aspects of HLLs' language abilities.

## 7. CONCLUSION

Collocations represent the conventionalized way of combining words in a given language, resulting in a statistical bond such that one word tends to predict the occurrence of the other. The current study focused on verbal collocations such as *cumplir una promesa* and *guardar silencio*. How do Spanish speakers come to know that the verb *cumplir* pairs with *promesa* but *guardar* does not? Researchers agree that collocations are a dimension of lexical knowledge that depends on massive exposure to input: "Knowledge about which words collocate with what is something that can only be learned from observing usage" (Dąbrowska, 2019, pg. 90). Furthermore, knowledge of

collocations develops in contexts of social and cultural integration, that is, when we observe the usage of speakers with whom we identify (Burdelski & Cook, 2012). These observations make collocations a top priority for the field of heritage language research. The current study represents the first step in understanding the collocational knowledge of Spanish HLLs by testing both form recognition and form recall. The results indicate that HLLs have robust knowledge of collocations in terms of recognition, but that production (recall) of collocations is more variable. Individual differences were evident, with significant correlations between language dominance and performance on both tasks. Finally, the participants' personal engagement with Spanish in daily activities (e.g., text messaging) accounts for some of the variance in performance.

## NOTES

1. Robles-Sáez's (2011) dictionary makes reference to both raw frequency and MI scores as criteria for inclusion: "En este libro hemos incluido las expresiones con frecuencia combinatoria alta y las de frecuencia combinatoria media-alta con una frecuencia de uso alta" (xi).
2. The confidence ratings are not analyzed in the current study.

## REFERENCES

- Birdsong, D., Gertken, L.M., & Amengual, M. (2012). *Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism*. COERLL, University of Texas at Austin. Web. <https://sites.la.utexas.edu/bilingual/>
- Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the effects of exercises on verb–noun collocations. *Language Teaching Research*, 18(1), 54–74.
- Boers, F., & Webb, S. (2018). Teaching and learning collocation in adult second and foreign language learning. *Language Teaching*, 51(1), 77–89.
- Burdelski, M., & Cook, H. M. (2012). Formulaic language in language socialization. *Annual Review of Applied Linguistics*, 32, 173–188.
- Carreira, M., and Potowski, K. (2011). Commentary: pedagogical implications of experimental SNS research. *Heritage Language Journal*, 8, 134–155.

- Dąbrowska, E. (2019). Experience, aptitude, and individual differences in linguistic attainment: A comparison of native and nonnative speakers. *Language Learning*, 69, 72–100.
- Davies, Mark. (2016–) *Corpus del Español: Two billion words, 21 countries*. <http://www.corpusdelespanol.org/web-dial/>
- Durrant, P. (2014). Corpus frequency and second language learners' knowledge of collocations: A meta-analysis. *International Journal of Corpus Linguistics*, 19(4), 443–477.
- Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics in Language Teaching*, 47(2), 157–177.
- Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research*, 26(2), 163–188.
- Ellis, N.C. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 33–68). Cambridge University Press.
- Fairclough, M. (2011). Testing the lexical recognition task with Spanish/English bilinguals in the United States. *Language Testing*, 28(2), 273–297.
- Fairclough, M. (2013). Muestras de lexicalización en estudiantes hispanos bilingües en los Estados Unidos: Un estudio exploratorio. *Revista Internacional d'Humanitats*, 27, 105–114.
- Fairclough, M., & Garza, A. (2018). The lexicon of Spanish heritage language speakers. In K. Potowski (Ed.), *The Routledge Handbook of Spanish as a Heritage Language* (pp. 178–189). Routledge.
- Foster, P. (2009). Lexical diversity and native-like selection: The bonus of studying abroad. In B. Richards et al. (Eds.), *Vocabulary studies in first and second language acquisition* (pp. 91–106). Palgrave Macmillan.
- Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, 67(S1), 155–179.
- Gatti, A., & O'Neill, T. (2017). Who are heritage writers? Language experiences and writing proficiency. *Foreign Language Annals*, 50(4), 734–753.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Allyn & Bacon.
- Gonnerman, L. M., Seidenberg, M. S., & Andersen, E. S. (2007). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General*, 136(2), 323–345.
- González Fernández, B., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have? *ITL-International Journal of Applied Linguistics*, 166(1), 94–126.

- González-Fernández, B., & Schmitt, N. (2019). Word knowledge: Exploring the relationships and order of acquisition of vocabulary knowledge components. *Applied Linguistics*, 41(4), 481–505. <https://doi.org/10.1093/applin/amy057>
- Gyllstad, Schmitt (2019). Testing formulaic language. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding Formulaic Language: A Second Language Acquisition Perspective* (pp. 174–191). Routledge.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
- Jensen, E. C. (2017). No silver bullet: L2 collocation instruction in an advanced Spanish classroom. *L2 Journal*, 9(3), 1–20.
- Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61(2), 647–672.
- Leal, T. (2018). Data analysis and sampling: Methodological issues concerning proficiency in SLA research. In A. Gudmestad & A. Edmonds (Eds.), *Critical reflections on data in second language acquisition* (pp. 63–88). John Benjamins.
- Macis, M., & Schmitt, N. (2017). Not just 'small potatoes': Knowledge of the idiomatic meanings of collocations. *Language Teaching Research*, 21(3), 321–340.
- Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research*, 21(3), 298–320.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223–242.
- Ortega, L. (2020). The study of heritage language development from a bilingualism and social justice perspective. *Language Learning*, 70, 15–53.
- Ortigosa, A. (2010). Convergencia lingüística en los calcos fraseológicos: Innovación estructural y semántica. *International Journal of the Sociology of Language*, 203, 27–44.
- Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research*, 20(1), 113–138.
- Pulido, M. F., & Dussias, P. E. (2020). Desirable difficulties while learning collocations in a second language: Conditions that induce L1 interference improve learning. *Bilingualism: Language and Cognition*, 23(3), 652–667.
- Robles-Sáez, A. (2010). *3000 locuciones verbales y combinaciones frecuentes*. Georgetown University Press.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.
- Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review*, 64(3), 429–458.
- Sonbul, S. (2015). Fatal mistake, awful mistake, or extreme mistake? Frequency effects on off-line/on-line collocational processing. *Bilingualism: Language and Cognition*, 18(3), 419–437.

- Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning*, 63(1), 121–159.
- Stengers, H., & Boers, F. (2015). Exercises on collocations: A comparison of trial-and-error and exemplar-guided procedures. *Journal of Spanish Language Teaching*, 2(2), 152–164.
- Szudarski, P., & Carter, R. (2016). The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics*, 26(2), 245–265.
- Toomer, M., & Elgort, I. (2019). The development of implicit and explicit knowledge of collocations: A conceptual replication and extension of Sonbul and Schmitt (2013). *Language Learning*, 69(2), 405–439.
- Treffers-Daller, J., Daller, M., Furman, R., & Rothman, J. (2016). Ultimate attainment in the use of collocations among heritage speakers of Turkish in Germany and Turkish–German returnees. *Bilingualism: Language and Cognition*, 19(3), 504–519.
- Van Osch, B., Hulk, A., Aalberse, S., and Sleeman, P. (2018). Implicit and explicit knowledge of a multiple interface phenomenon: Differential task effects in heritage speakers and L2 speakers of Spanish in the Netherlands. *Languages*, 3(3), 25.
- Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics*, 32(4), 430–449.
- Wolter, B., & Yamashita, J. (2015). Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics*, 36(5), 1193–1221.
- Yamashita, J., & Jiang, N. A. N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly*, 44(4), 647–668.
- Zyzik, E. (2021). El conocimiento léxico de los hablantes del español como lengua de herencia. In J. Torres and D. Pascual y Cabo (Eds.), *Aproximaciones al estudio del español como lengua de herencia*. Routledge.

**APPENDIX**

Collocation	Raw Frequency	Mutual Information (MI)	Rank frequency of component words
colmar la paciencia	303	8.12	colmar: 7246; paciencia: 2366
guardar las apariencias	324	4.09	guardar: 974; apariencia: 2524
cerrar la llave	376	6.12	cerrar: 758; llave: 3141
aceptar las disculpas	431	3.11	aceptar: 461; disculpa: 3822
correr con los gastos	521	3.88	correr: 668; gasto: 1801
estirar las piernas	545	8.91	estirar: 5398; pierna: 1884
mantener bajo control	621	6.34	mantener: 213; control: 450
vencer el miedo	965	5.38	vencer: 1823; miedo: 553
dejar plantado	970	4.11	dejar: 88; plantado: 3427
pasarse de la raya	1112	8.9	pasar: 83; raya: 5249
ceder el paso	1163	4.64	ceder: 2598; paso: 277
romper las reglas	1185	5.44	romper: 1083; regla: 1001
atar cabos	1187	12.84	atar: 3440; cabo: 616
hacer las paces	1214	6.04	hacer: 21; paz: 634
leer entre líneas	1257	8.46	leer: 169; línea: 425
llevar ventaja	1282	3.75	llevar: 98; ventaja: 1153
morderse la lengua	1282	8	morder: 4559; lengua: 1175
levantar cabeza	1328	4.94	levantar: 880; cabeza: 489
saldar una deuda	1344	10.11	saldar: 8812; deuda: 1079
partir de cero	1370	3.1	partir: 904; cero: 1960
aprobar un examen	1383	7.05	aprobar: 1063; examen: 1486
mover un dedo	1557	6.82	mover: 695; dedo: 1600
montar un negocio	1635	7	montar: 1996; negocio: 414
inspirar confianza	1688	6.99	inspirar: 2031; confianza: 1172
guardar un secreto	1861	6.42	guardar: 974; secreto: 1830
romper el corazón	2012	3.8	romper: 1083; corazón: 427
marcar un gol	2093	7.77	marcar: 824; gol: 1514
llegar lejos	2160	4.19	llegar: 90; lejos: 941
levantar la voz	2430	4.94	levantar: 880; voz: 530
entrar en detalles	2441	9.78	entrar: 319; detalle: 795
hacer cola	2575	4.14	hacer: 21; cola: 3105

(Appendix Continued)

(Appendix Continued)

Collocation	Raw Frequency	Mutual Information (MI)	Rank frequency of component words
correr el peligro	2887	7.92	correr: 668; peligro: 1140
tocar fondo	2976	4.23	tocar: 499; fondo: 1109
pasar de largo	2984	4.98	pasar: 83; largo: 392
cubrir los gastos	3053	7.79	cubrir: 992; gasto: 1801
cumplir una promesa	3093	6.29	cumplir: 391; promesa: 1913
volver a la normalidad	3155	5.9	volver: 178; normalidad: 4708
hacer el ridículo	3162	3.53	hacer: 21; ridículo: 5768
surtir efecto	3399	8.6	surtir: 8183; efecto: 612
violar la ley	3586	5.35	violar: 2189; ley: 309
recaudar fondos	3593	10.31	recaudar: 4586; fondo: 1109
dejar en paz	4191	4.57	dejar: 88; paz: 634
meter la pata	4208	8.38	meter: 683; pata: 2957
pasar hambre	4406	4.05	pasar: 83; hambre: 1581
tomar las riendas	4412	7.72	tomar: 129; rienda: 5737
sufrir las consecuencias	4697	5.78	sufrir: 477; consecuencia: 653
guardar silencio	5510	6.98	guardar: 974; silencio: 1350
cambiar de opinión	5758	6.55	cambiar: 223; opinión: 329
cometer un delito	6002	7.7	cometer: 1030; delito: 1068
rendir homenaje	7809	9.78	rendir: 2020; homenaje: 2657
ganarse la vida	8693	5.7	ganar: 284; vida: 56
tomar en serio	9322	6.54	tomar: 129; serio: 821
abrir camino	9685	3.18	abrir: 289; camino: 301
poner en duda	11618	7.47	poner: 86; duda: 558
poner a prueba	11884	7.56	poner: 86; prueba: 486
salvar la vida	12043	4.24	salvar: 1193; vida: 56
llegar a un acuerdo	15243	6.83	llegar: 90; acuerdo: 802
prestar atención	35147	8.46	prestar: 1101; atención: 368
tomar una decisión	49475	6.4	tomar: 129; decisión: 339
llamar la atención	72896	8.09	llamar: 151; atención: 368

**Note:** The rank frequency is given for the lemma. In cases where a collocation includes a preposition, frequency information for the preposition is not provided but assumed to be very high.