Spanish L1 EFL learners' recognition knowledge of English academic vocabulary: the role of cognateness, word frequency and length

Raquel Perez Urdaniz and Sophia Skoufaki*

Department of Language and Linguistics, University of Essex, Colchester, UK

Raquel Perez Urdaniz (address where the research project was conducted)

Department of Language and Linguistics

University of Essex

Wivenhoe Park

Colchester

CO4 3SQ

UK

Raquel Perez Urdaniz (present address)

C/Cendea de Ansoain, Nº13

Berriozar, 31013

Navarra

Spain

Email: raquelperezurdaniz@hotmail.com

ORCID ID: 0000-0001-7777-9794

Sophia Skoufaki

Department of Language and Linguistics

University of Essex

Wivenhoe Park

Colchester

CO4 3SQ

UK

Tel.: + 44 (0) 1206 873754

Email: sskouf@essex.ac.uk

ORCID ID: 0000-0002-9992-3977

Spanish L1 EFL learners' recognition knowledge of English academic vocabulary: the role of cognateness, word frequency and length

Abstract: Academic vocabulary knowledge predicts students' academic achievement across educational levels. English academic vocabulary knowledge is especially valuable because English is used in academia worldwide. Therefore, examining the factors that can predict English academic vocabulary knowledge can inform pedagogy, thus indirectly boosting students' chances of academic success around the world. This study examines the extent to which cognateness, word frequency and length predict the ability of English as a Foreign Language (EFL) learners who have Spanish as their first language (L1) to recognise written English academic words. 38 Spanish L1 university students' recognition knowledge of English cognates was measured via a Yes/No test containing words sampled from the most frequent 1,000 lemmas of the Academic Vocabulary List (Gardner and Davies 2014). 34 participants' data were retained in the final analysis, a multiple regression with item facility (IF) as the outcome variable and word frequency, cognateness and word length as predictors. Most of the IF variance is explained by word frequency, followed by cognateness and finally a frequency by cognateness interaction whereby word frequency is more predictive of IF for non-cognates than cognates. These findings indicate that academic cognate-word awareness raising activities can be worthwhile. Implications for research and pedagogy are discussed.

Keywords: cognate words, cognate facilitation, word frequency, word length, academic vocabulary, vocabulary learning, Spanish-English bilingualism

1 Introduction

Academic vocabulary is broadly defined as the vocabulary used in academic writing and speech without being specific to any discipline but opinions vary on which words these are exactly (see Paquot 2010: 9-28 for an overview). The various conceptualisations of *academic vocabulary* notwithstanding, research consistently indicates that at various educational levels academic vocabulary can predict the quality of students' performance in both productive and receptive tasks (e.g., Csomay and Prades 2018; Kieffer and DiFelice Box 2013; Truckenmiller and Petscher 2019) as well as their overall academic achievement (e.g., Schuth et al. 2017; Townsend et al. 2012).

Unfortunately, immersion in an educational setting does not sufficiently foster the learning of academic vocabulary. For example, English academic vocabulary knowledge develops incidentally very slowly for international university students studying in English-speaking universities. In Schmitt (1998) four international postgraduate students at a British university were interviewed about their knowledge of 11 academic words three times during an academic year. Results suggest that by the end of the year some aspects of vocabulary knowledge had been mastered (e.g., spelling) but not others (e.g., meaning senses of polysemous words, word family members). Moreover, even when academic vocabulary instruction interventions take place, academic vocabulary knowledge may still be acquired slowly, even by native speakers of English (e.g., Spencer et al. 2017). Research on the factors that affect English academic vocabulary learning can help to enhance the effectiveness and efficiency of English academic vocabulary instruction: finding out which lexical characteristics make words challenging to learn can help researchers recommend i) direct instruction for challenging words and ii) the provision of vocabulary awareness raising activities and instruction on vocabulary learning strategies so that learners can learn efficiently on their own words that are easy to learn. Such research is worthwhile given the role of English as the most widespread lingua franca and its dominance in academia worldwide (Melitz 2018).

Examining whether certain lexical characteristics make some academic words easier to learn than others could help identify words which lend themselves to teaching, thus paving the way for efficient academic vocabulary instruction. With this rationale in mind, Nation (1990) suggests prioritising the teaching of cognates over non-cognates because he expects that the form and meaning overlap between L1 words learners already know and second language (L2) target words will facilitate L2 word learning. Research on the role of three such lexical factors which can potentially predict English academic vocabulary knowledge, namely, cognateness, word frequency and length, has been conducted in relation to general English vocabulary learning. Given the lack of such research in relation to English academic vocabulary, the present study examines how predictive cognateness, word frequency and length are of EFL learners' ability to recognise English academic vocabulary.

In this study participants are Spanish L1 EFL learners. This L1-L2 pair was chosen due to a pedagogical consideration. Research suggests that Spanish-English bilinguals do not recognise all cognates as such (e.g., Nagy et al. 1993; August et al. 2005). Therefore, if the present study indicates that cognateness significantly predicts academic word recognition knowledge, awareness raising activities about academic cognates are likely to be useful in primary, secondary and higher EFL education for Spanish L1 learners. Moreover, since a large proportion of English academic words are Spanish cognates

(Lubliner and Hiebert 2011: 80), these awareness raising activities can foster considerable academic vocabulary learning gains, which can, in turn, help improve Spanish EFL students' performance in educational settings where English is required.

2 Literature review

2.1 The role of cognateness in (English) L2 vocabulary learning

Cognates have been defined mainly from two viewpoints (Daulton 2008). From a diachronic viewpoint, cognates are etymologically related words similar in meaning and form. From a synchronic viewpoint, cognates are words which, irrespective of whether they are etymologically related or not, are similar in form and meaning. The synchronic definition is the most relevant to language learning research because language learners' *perceived* cross-linguistic similarities, not historical linguistics analyses, play a role in language learners' recognition of a word as a cognate (Daulton 2008).

A word which has a cognate that a language learner knows already in another language tends to be learned more efficiently or processed (e.g., uttered in a picture naming task, recognised as a real word in a lexical decision task) faster than a word for which there is no such cognate or the language learner does not know this cognate (Petrescu et al. 2017). A possible reason for this phenomenon, called *cognate facilitation*, is learners' presumption that words which are phonologically and/or orthographically similar to words they already know are likely to have similar meanings (Daulton 2008; Jiang 2000 but see Kellerman 1978). Research findings which suggest that the degree to which two cognates resemble each other modulates the cognate facilitation effect (e.g., Allen and Conklin 2013; Otwinowska and Swewczyk 2017) seem congruent with this claim. Another possible reason is that, according to some models of vocabulary learning (e.g., Ecke and Hall 2014; Jiang 2004), at initial stages of vocabulary learning a new word shares a mental-lexicon entry with the most similar phonologically or orthographically known word whereas non-cognate words require new language-specific mental lexicon entries.

The facilitative effect of cognates in the learning of L2 English words has been suggested in language testing and language learning research. A few studies on language testing explored the reasons behind Differential Item Functioning (DIF) between EFL learner groups with different L1s; these studies will be reviewed here. DIF occurs when one learner group has a significantly lower average score for some test items than another learner group of similar proficiency level (Zumbo 1999). Chen and Henning (1985) investigated language bias in an English as a second language (ESL) placement test by comparing the response patterns of L1 Spanish and L1 Chinese learners. Chen and Henning (1985) expected to find positive bias for the L1 Spanish learners because they considered English and Spanish as linguistically closer than English and Chinese; they based this assumption on the fact that English and Spanish are Indo-European languages whereas Chinese is part of the Sino-Tibetan language family. Four test items indicated bias, all of them favouring L1 Spanish learners because they were Spanish cognates. Sasaki (1991) compared two techniques used for the identification of DIF in an ESL placement test for L1 Spanish and L1 Chinese learners, one technique being the same as that used in Chen and Henning (1985). All the test items which exhibited DIF included cognate words which favoured L1 Spanish test-takers.

The role of L1-cognate word frequency on DIF was examined in Stoeckel and Bennett (2013) and Bennett and Stoeckel (2014a, 2014b). In Stoeckel and Bennett (2013), in a vocabulary test given to Korean and Japanese learners of EFL, 21 items showed DIF. Out of these DIF items, 10 were loanwords either in Korean or in Japanese. The frequency of the tested word's cognate in the test-takers' native language predicted the direction of DIF for all except two of the loanwords. Bennett and Stoeckel (2014a) also suggests that L1 loanword frequency predicts the direction of DIF; infrequent L1 loanwords may not be a source of DIF, whereas high frequency L1 loanwords are useful predictors of DIF. Bennett and Stoeckel (2014b) examined the effect of English-word frequency and Japanese cognate-word frequency on the English vocabulary test scores of Japanese college students. Results indicated that Japanese-word frequency was a better predictor of English word knowledge than English-word frequency.

The role of target-word cognateness in Vocabulary Size Test (VST) (Nation and Beglar 2007) scores was examined in Beglar (2010) and Elgort (2013). In Beglar (2010) L1 Japanese EFL learners' scores for VST items from a number of British National Corpus (BNC) frequency bands (namely, the second to the fourth and the eighth) were higher than expected based on vocabulary corpus frequencies. Beglar (2010: 109) attributed this finding to the large number of Japanese loanwords. Elgort (2013) compared the results of the monolingual and the bilingual versions of the VST in the same L1 Russian intermediate-level EFL learners. Test responses were on average more accurate for cognates than for noncognates. A 'cognateness by frequency' interaction approached significance. This interaction effect was interpreted as the moderation of the effect of cognateness on accuracy by frequency. This moderation effect was attributed to the fact that whereas intermediate-level learners know receptively many high frequency words, irrespective of whether they are L1 cognates or not, they know receptively more low-frequency cognates than non-cognates.

Laufer and McLean (2016) examined the effect of cognateness, L1, proficiency level and vocabulary test type on vocabulary test scores. The same participants' scores were compared between versions with and without cognate words of a form recall, a meaning recall and a form recognition test. Their participants were beginner, intermediate, and advanced EFL learners whose L1 was Hebrew or Japanese. Due to low Japanese L1 participant numbers, the average scores of their test results could not be compared across tests. The results of the Hebrew L1 participants indicate that for all tests the cognates version received higher scores than the version without cognates. A comparison of scores in each test across beginner, intermediate and advanced EFL learners suggests that scores were higher for cognate than for non-cognate items only for beginner and intermediate learners in the recall tests; they were higher for cognate than non-cognate items for all proficiency groups in the recognition tests. These findings suggest that cognate facilitation depends on both the learners' proficiency level and the kind of vocabulary test administered.

The role of cognateness in vocabulary learning has also been examined vis a vis the role of word frequency and length. Before reviewing the few studies which examined the extent to which all these three factors can predict L2 vocabulary knowledge, we will briefly review research on whether word frequency and length predict vocabulary knowledge.

2.2 The role of word frequency in English vocabulary learning

In most psycholinguistic and applied linguistic research, word frequency has been operationalised as the number of occurrences of the target lexical items in a large reference corpus, such as the BNC or COCA (e.g., Horst et al. 1998; Reynolds et al. 2015). Some applied linguistics studies operationalised word frequency (also) as the frequency of occurrence of a word in a learner's classroom input or instruction materials (e.g., Demetriou 2017; Vidal 2003) and in the input they receive from other sources (e.g., reading outside the classroom, watching videos) (e.g., Horst et al. 1998; Peters and Webb 2018) because a frequent word in a native speaker corpus is not necessarily encountered frequently by a foreign language learner. Although both studies using a corpus-based and those using an L2-input based definition of *word frequency* vary in whether word frequency is that of a word form (whereby the frequencies of distinct word forms, such as *learn* and *learns,* will be counted separately), lemma (whereby the frequencies of the root form and of all its inflected forms will be added up), or word family (whereby the frequencies of the root form, all its inflected forms and any derived forms will be added up), most studies indicate that the more frequent a word is, the more likely it is to be learned (Reynolds and Wible 2014). Studies which do not suggest a positive relationship between lexical frequency and lexical learning (e.g., Macis and Schmitt 2017; Pellicer-Sánchez 2017; Szudarski and Carter 2016) indicate that vocabulary learning is the result of the interplay of various factors, some of which may override or moderate the role of lexical frequency. Hence, an examination of the role of academic vocabulary frequency on academic vocabulary learning when other factors, such as cognateness, are taken into consideration, is warranted.

2.3 The role of word length in English vocabulary learning

The role of word length in English word learning has been examined in a few studies. Laufer's (1990, 1997) reviews of relevant research indicate that although most studies suggest that the longer a word is, the more difficult it is to learn, a few studies failed to find this effect. In these studies word length was measured as the number of syllables in a word. Perhaps the use of this measure and/or other methodological aspects of these studies – Laufer (1990) indicates that not all studies distinguished the effect of word length from that of other lexical characteristics on word learning – led to these conflicting results among studies.

In addition to measuring word length as the number of syllables per word, recent relevant studies measured it as the number of letters and the number of phonemes per word. Culligan (2015) tested the same language learners via three vocabulary tests, namely, a Yes/No test, the Vocabulary Knowledge Scale (Paribakht and Wesche 1997) and the Vocabulary Levels Test (Nation 1983), and then correlated these test scores with word length measured in three ways (the number of letters per word, the number of phonemes per word, and the number of syllables per word) and word frequency measured in five ways, according to five English corpora. Correlations with test scores were moderate for both the word frequency and length measures. This finding suggests that word length and word frequency cannot predict well the word knowledge measured by the three vocabulary knowledge tests used in Culligan (2015). Alsaif and Milton (2012) aimed to characterise the vocabulary included in EFL textbooks used in secondary schools in Saudi Arabia. It also examined the relative extent to which a word's frequency of occurrence in the textbooks, length measured as the number of syllables in a word and concreteness predicted male secondary public school pupils' knowledge of a sample of words that appeared in these textbooks. All these factors predicted vocabulary knowledge significantly, with word length being the most important factor.

2.4 The relative role of cognateness, word frequency and length in (English) vocabulary learning

The extent to which L2-cognate word frequency, L2 word length (operationalised as the number of syllables in each word) and cognateness (operationalised as the letters that the L1 and the L2 words shared) predict British learners' French lexical recognition knowledge was investigated in Milton and Daller (2007). Participants' proficiency levels in French ranged from beginner to 'degree level' (Milton and Daller 2007, presentation slide 8). French words were tested via a Yes/No test, which consisted of 100 items selected from the 5,000 most frequent words from a lemmatised French wordlist. Multiple regression analysis suggested that only frequency significantly predicted word recognition knowledge. Milton (2009: 41) attributed the lack of a significant effect for cognateness and word length to various reasons. First, if cognateness and word length had been operationalised differently, significant effects might have been found for them. This explanation is plausible because significant effects were indeed found in Willis and Ohashi (2012), where cognateness was operationalised as a categorical instead of a continuous variable and word length was operationalised as the number of phonemes per word[1]. Second, French and English have so many cognate words that the cognate effect might have been hidden. Finally, thanks to similarities between English and French morphology (e.g., they have many similar affixes), participants could have inferred the meaning of unknown long words by dividing them into their component parts.

[1] Willis and Ohashi (2012) will be reviewed in the next paragraph.

Milton and Daller (2007) was partially replicated in Willis and Ohashi (2012). The extent to which cognateness, word length and word frequency can predict word knowledge was examined but participants were Japanese EFL learners with 'a wide range of proficiency levels' (Willis and Ohashi 2012: 128) and vocabulary size was measured via part of the VST instead of a Yes/No vocabulary test. Moreover, cognateness was operationalised as a binary instead of a gradable concept. Contrary to Milton and Daller (2007), cognateness was the most important predictor of word learnability, followed by frequency and word length measured as number of phonemes per word. These results suggested that cognates are easier to learn and retain than non-cognates and that their item facility (IF)[2] is higher than the IF of non-cognates with the same frequency and length.

Apart from examining the role of the characteristics examined in Milton and Daller (2007) and Willis and Ohashi (2012), Reynolds et al. (2015) examined the role of these factors on IF: a) the number of word family members in the family of a tested word (e.g., according to Nation's 2012 BNC word family lists, *poor* belongs to a word family of five, *poor*, *poorer*, *poorly*, and *poorness*), b) whether the tested word was polysemous or not according to WordNet (Fellbaum 1998) and c) whether the tested word was a noun, verb or adjective. Participants were advanced Taiwanese EFL learners. They were tested via the VST, as in Willis and Ohashi (2012). The only statistically significant predictors of IF were polysemy, frequency and whether a word is a noun or not. Reynolds et al. (2015: 139) attribute the non-significant effect of cognateness on IF to the limited number of English-Chinese cognates in the VST; they call for the use of other vocabulary tests to

---

[2] IF is a statistic used to analyse the percentage of participants who correctly answer a given item. It is calculated by adding up the number of participants correctly answering an item and dividing it by the total number of participants (Brown 2005).

examine the role of cognateness in English vocabulary learning by native speakers of Chinese.

Reynolds et al. (2018) examined the role of polysemy, part of speech, word length, word family size and frequency of exposure to words on IF. Although this study did not examine the role of cognateness on Taiwanese learners' vocabulary acquisition, it is briefly reviewed here because it is similar in design and aims to Reynolds et al. (2015). Frequency of exposure to words was operationalised as a) lemma frequency in the BNC and b) the inclusion of words in the College Entrance Examination Center's Reference Word List (RWL) (CEEC 2002). Participants' vocabulary knowledge was tested via the VST. RWL level and inclusion of a word in the RWL explained most of the IF variance. This finding suggests that frequency of exposure to a word is a more important factor than frequency of occurrence in a large native English corpus and the other lexical-characteristic predictors. As in Willis and Ohashi (2012), word length operationalised as the number of phonemes per word had a moderate effect on IF.

The studies reviewed so far cannot offer evidence of a causal link between cognateness and frequency on the one hand and vocabulary learning on the other because they used covariance-based statistical analyses. However, the experimental studies that will be reviewed in this paragraph indicate that cognateness and frequency significantly affect vocabulary learning. In Lotto and De Groot (1998) target Italian words and their L1_equivalents were presented in a paired-associate task to native speakers of Dutch who had no knowledge of Italian. Form recall for the Italian words was compared between a) words cognate and non-cognate to Dutch words and b) words which were the translation equivalents of high- versus low-frequency Dutch words. Significantly more cognate words were recalled correctly than non-cognate words in all frequency levels. These findings were replicated in De Groot and Keijzer (2000), which modified the method of

Lotto and de Groot (1998) by teaching participants pseudocognates instead of real words. Finally, in two experiments with Dutch-speaking EFL learners Peters and Webb (2018) found that watching a one-hour documentary led to significant vocabulary gains in terms of meaning recall and recognition and that vocabulary learning was affected by frequency of occurrence, prior vocabulary knowledge and cognateness.

2.5 The role of cognateness in English academic vocabulary learning

Although the role of cognateness in English vocabulary learning in general has been examined extensively, to our knowledge, its role in academic vocabulary learning has been examined in only two studies. Daulton (2005) compared the scores of L1 Japanese first-year undergraduate students in a revised version of the Vocabulary Levels Test (VLT) (Schmitt 2000). This version of the VLT contains words from the two most frequent 1,000 word bands and academic words from the Academic Word List (AWL) (Coxhead 2000). High scores were achieved in both the general-vocabulary and academic-vocabulary sections of the test. They were attributed to the high percentage of cognates among the items (Daulton 2005: 5).

Petrescu et al. (2017) compared the scores of Romanian L1 and Vietnamese L1 ESL university students across the general, academic and rarer vocabulary sections of a modified version of the VLT (Schmitt 2000). Since both student groups' scores were very high in the academic vocabulary subtest, results suggest a significant but moderate cognate facilitation effect on receptive academic vocabulary knowledge. However, this near ceiling effect may be an artefact of the test, which included only words sampled from the AWL, an academic vocabulary list which largely consists of high-frequency words (see section 2.6).

Research on the role of cognateness in EFL academic vocabulary learning has important consequences for teaching given the preponderance of academic word cognates in certain L1s. For example, Study 2 in Daulton (2005) explored the correspondences between academic English words, operationalised as the AWL, and common loanwords in Japanese. A quarter of the AWL corresponds to frequent Japanese words.

According to Lubliner and Hiebert (2011), Spanish-speaking EFL learners may have a similar advantage when learning English. Spanish and English have around 10,000 Latin-based cognate word lemmas in addition to cognates of other origins (Nash 1997). However, a positive effect of cognateness on English academic vocabulary learning for Spanish L1 speakers should not be taken for granted because Spanish L1 speakers do not always recognise English cognates as such (August et al. 2005; Nagy et al. 1993). Why cognate recognition is sometimes problematic is still unclear; it has been suggested that cognate transparency is influenced by various factors such as individual differences among learners, exposure to cognate instruction and semantic, phonological and spelling dissimilarities in cognate pairs (August et al. 2005).

2.6 Operationalising English academic vocabulary

The traditional conceptualisation of academic vocabulary is based on the assumption that it is rarer than the most frequent words (e.g., Nation 1990). The AWL embodies this conceptualisation because the starting point for its compilation was the decision to exclude from it which appear in the General Service List (GSL) (West 1953), a list of the most common 2,000 English word families (Coxhead 2000: 218). Research comparing the AWL and wordlists containing high-frequency words other than the GSL, such as wordlists derived from the British National Corpus (Cobb 2010; Masrai and Milton 2018), indicates a large overlap between the AWL and frequent words. The fact that the AWL is imbalanced across academic disciplines (e.g., Hyland and Tse 2007) also indicates that a wordlist that contains words evenly distributed across the academic subsections of a corpus, is more likely to include vocabulary that is used across disciplines.

The Academic Vocabulary List (AVL) (Gardner and Davies 2014) consists of words from all frequency bands of the Corpus of Contemporary American English (COCA). The AVL consists of 3,014 word lemmas[3] which occur at least 50% more frequently in the Academic section of COCA than would normally be expected, are evenly distributed across disciplinary sections of the Academic section of COCA and occur in at least seven of COCA's nine disciplinary sections. In this study academic vocabulary was operationalised as the AVL because the assumed distinction between frequent and academic vocabulary has been contested (Cobb 2010; Masrai and Milton 2018) and because the AVL was constructed in a more methodologically stringent way than the AWL (Gardner and Davies 2014: 312-6).

3 The present study

[3] The Excel file with the AVL provided as supplementary material in Gardner and Davies (2014) consists of 3,105 lemmas but the entry for the word *disproportionately* appears twice (Durrant, 2016: 53). Therefore, the real number of lemmas in the AVL is 3,014.

The present study examines how well lexical characteristics predict *academic* rather than general vocabulary knowledge. This research aim is warranted not only because such research has been sparse but also because it may inform EFL and English for Academic Purposes (EAP) pedagogy by guiding the selection of academic vocabulary that is easy to teach. The development of academic vocabulary knowledge in turn can help learners improve their performance in academic tasks. For example, it can help them approach the 98% text coverage necessary for good reading comprehension (e.g., Schmitt et al. 2011) and meet the requirements of university-level academic writing (Durrant 2016).

Focusing on the role of cognateness in the learning of EFL academic vocabulary by Spanish L1 learners is worthwhile because no study has examined whether Spanish cognateness significantly predicts English academic vocabulary knowledge despite the high proportion of Spanish cognates in English academic vocabulary; 74.7% of the AWL headwords are Spanish cognates (Lubliner and Hiebert 2011: 80)[4].

Finally, this study is warranted because it avoids some of the methodological shortcomings of previous studies. Cognateness and word frequency were confounded in earlier studies. By testing English vocabulary knowledge via the VST (Nation and Beglar 2007), Willis and Ohashi (2012) tested cognate words which appeared among words sampled from the seven most frequent 1,000 word-family bands in the spoken section of the BNC. The number of Japanese-English cognates tested varied among frequency bands (e.g., four cognate words are in the first band whereas only two cognate words are in the third band). Due to this uneven distribution of cognates among frequency bands, cognateness and frequency are not clearly distinct variables in the VST. Consequently, in the multiple regression analysis reported in Willis and Ohashi (2012) some of the IF variance actually due to cognate words' frequency may have been attributed wrongly to

[4] To our knowledge, no study has estimated the percentage of cognate words in the AVL.

the fact that they were cognates. Cognates and non-cognates were not matched for frequency in Petrescu et al. (2017), a study which examined cognate facilitation in academic vocabulary learning. Although Yes/No tests are not without their shortcomings (see section 7), the use of a Yes/No test enables us to examine with greater precision the role of word frequency on the learning of English cognate words in the present study. An equal number of cognate and non-cognate words were sampled from the same wordlist frequency band to examine the unique contribution of frequency and cognateness towards academic word recognition knowledge.

Using a Yes/No test instead of the VST also enabled us to have a higher sampling rate than some earlier studies. The sampling rate in the VST (which was used in Willis and Ohashi 2012, Reynolds et al. 2015 and Reynolds et al. 2018) is 10 words per 1,000 word families band (Nation and Beglar 2007). According to Gyllstad et al. (2015) a sampling rate of 10 items per 1,000 in the VST is not sufficient for the findings to be generalisable to all 1,000 items while higher sampling rates (they tested the reliability of up to 30 items sampled from 1,000) lead to satisfactory levels of reliability. In our study, not only were test items selected through stratified sampling (see section 4.2.2) but also 52 lemmas (26 cognates and 26 non-cognates) from the most frequent 1,000 AVL lemmas were tested to increase the generalisability of our findings to the whole list of the 1,000 most frequent AVL lemmas.

In sum, the present study addresses these questions:

(1) Which combination of L2 word variables (cognateness, frequency, length) best predicts the L2-English recognition academic vocabulary knowledge of L1 Spanish university students?

(2) What is the contribution of each of the L2 word variables examined in this study to L1 Spanish university students' L2-English recognition academic vocabulary knowledge?

4 Methodology

This study examined the effect of cognateness, L2 frequency and word length, on the learning of English academic words by Spanish university students. Participants sat a Yes/No test which contained English academic-word cognates and non-cognates matched for frequency. The aforementioned lexical characteristics were measured and used as predictors in a regression analysis (see Table 1 for descriptive statistics of the predictors and the outcome variable). The operationalisation of cognateness, word frequency and length will be summarised below.

Cognate identification approaches depend on the researcher's working definition of cognateness (Potapova et al. 2016). Because different approaches can lead researchers to different categorisations of words in terms of cognateness (Potapova et al. 2016), we combined identification approaches as a means of triangulation. We first categorised words on the basis of a English-Spanish cognates database (Montelongo et al. 2011) and then asked English-Spanish bilinguals to categorise these words as cognates and non-cognates. Both approaches are congruent with our working definition of cognates as words which are similar in terms of their form and meaning in two languages (see Otwinowska and Szewczyk 2017, Peters and Webb 2018 for very similar working definitions). In practical terms, first a candidate word for the Yes/No test was considered a cognate if it appeared in the Find-A-Cognate database (Montelongo et al. 2011). After shortlisting words according to various considerations (see section 4.2.2), the final list of cognate candidate words and their non-cognate frequency-matched words were categorised by four fluent speakers of English and Spanish between cognates and non-cognates. Section 4.2.2 summarises the procedure followed to create the items list for the English Yes/No test in detail.

Word form frequency was measured according to the SUBTLEX-UK word frequency database (van Heuven et al. 2014). The SUBTLEX-UK database provides subtitle-based UK frequency norms from television channels directed to children. This word frequency database was used because British English is the variety of English taught in Spain. SUBTLEX is a family of lexical databases based on corpora of film and TV programmes in various languages. A SUBTLEX database was preferred over another database or frequencies from another corpus (e.g., BNC) because word frequencies in film and television approximate those people are exposed to through social interaction better than word frequencies from written texts or spoken and written texts taken together (Brysbeart and New 2009: 979).

This study also tested participants' a) vocabulary size through a standardised test, LexTale (Lemhöfer and Broersma 2012), to estimate their EFL proficiency level (see section 4.2.1) and b) knowledge of the Spanish words which were equivalent to the L2 target cognate words to establish whether any of the Spanish cognate words were unknown to them (see section 4.2.3) so as to exclude such L2 cognate words from data analysis.

4.1 Participants

Participants were 38 Spanish university students who were studying various subjects at two universities in Spain. A biodata and language learning history questionnaire asked students about their personal details (gender, age, mother tongue, research discipline), for how many years they had been learning English and how long they have stayed in an English-speaking country, if so. This questionnaire appears in Appendix C.

All participants had studied English formally for at least four years. Their age ranged from 18 to 32 (Mean = 21.11, SD = 2.73). 28 (73.68%) of participants were in their 20s; 9 (23.68%) were 18 or 19 years old. One participant was 32 years old. All participants were Spanish native speakers; 13 reported being native speakers of Basque as well. 22 had stayed in an English-speaking country. Most of them had stayed there for less than a month (six participants), between one and two months (six participants) or between three to six months (five participants). Two participants had stayed for 10 months and three for a year.

LexTALE results indicate 18 of them were advanced (i.e., at the C1 or C2 CEFR levels), 19 were upper-intermediate level learners (i.e., at the B2 CEFR level) and one was a low-intermediate learner (i.e., at the B1 CEFR level) (see section 4.2.1 for information on LexTALE).

4.2 Data collection instruments

The online vocabulary size test LexTALE (Lemhöfer and Broersma 2012), an English vocabulary test and a Spanish vocabulary test were the data collection instruments in this study. All of them are Yes/No tests. In Yes/No tests test-takers are given a list of words and they indicate the words they know by ticking them (in pencil-and-paper tasks) or selecting them (in online tasks). These tests have been used for various research purposes in several studies (e.g., Eyckmans 2004; Masrai and Milton 2018; Milton and Daller

2007). Research suggests that Yes/No tests are a reliable way of measuring L2 learners' recognition vocabulary size (e.g., Huibregtse et al. 2002) although, as all tests, they are not without their limitations (see section 7).

4.2.1 LexTALE

The vocabulary test LexTALE[5] (Lemhöfer and Broersma 2012) provided an estimate of participants' overall English proficiency level. LexTALE was considered a good indicator of English proficiency level because its scores correlate significantly with scores in two English proficiency tests, the Quick Placement Test (QPT) and the Test of English for International Communication (TOEIC) (Lemhöfer and Broersma 2012). Based on equivalences between LexTale score bands and QPT score bands on the one hand and QPT score bands and Common European Framework of Reference for Languages (CEFR) levels on the other, Lemhöfer and Broersma (2012: 341) suggest that LexTALE scores below 59% suggest lower-intermediate (B1) or lower proficiency level, those between 60 and 80% suggest upper-intermediate (B2) proficiency level and those between 80 and 100% suggest advanced (C1 and C2) proficiency level. LexTALE has the additional advantage over other vocabulary tests of requiring only 3-5 minutes to complete.

4.2.2 English Yes/No vocabulary test

[5] This computerised vocabulary test can be downloaded or sat online at http://www.lextale.com.

The main data collection instrument was an English Yes/No vocabulary test administered to examine the role of three lexical variables (cognateness, frequency and length) in academic English word learnability. A Yes/No test was preferred over a multiple-choice recognition test (e.g., see Willis and Ohashi 2012) because it enables one to test many lexical items fast (Culligan 2015; Huibregtse et al. 2002).

This test consisted of 87 items; 26 were English words which had Spanish cognates, 26 were English words without Spanish cognates, and 35 were pseudowords. Pseudowords were used to limit the participants' tendency to select as known words which they do not really know. Since the ideal proportion of words to pseudowords is unknown (Eyckmans 2004), we followed Meara and Buxton's (1987) suggestion for pseudowords to equal 66.66% of the number of words. All test items appear in Appendix A.

The academic English cognate words were selected from a stratified sample of the 1,000 most frequent lemmas in the AVL (Gardner and Davies 2014). Test items were not samples from the other two AVL 1,000 lemma bands because we were afraid that if the English Yes/No test had triple the length (that is, 261 test items) participants would have been tempted to skip items so as to finish the test fast. We consider this eventuality likely because participants also had to do two other vocabulary tests – LexTale and the Spanish Yes/No test – and fill in a biodata questionnaire.

To create this stratified sample, the frequency-ordered list of the 1,000 most frequent AVL lemmas was first divided into 100-lemma sections. Then each section's percentage breakdown in terms of Part of Speech (POS) was calculated. 10 lemmas were selected from each of the 100-lemma sections of the list with the aim to preserve each 100-lemma section's POS percentage breakdown in the 10-lemma sample. For example, the POS percentage breakdown in the first 100 AVL lemmas was 66 nouns, 14 adjectives, 17 verbs, and 3 adverbs. 7 nouns, 1 adjective, and 2 verbs were randomly selected to create a 10-lemma sample that would reflect as much as possible the POS breakdown in this 100-lemma section.

The 100 lemmas sampled from the AVL following the procedure summarised above were looked up in the dictionary function of Lexico, the online database of English words by Oxford University Press, to see whether each could be categorised as a member of only one POS[6]; lemmas ambiguous between more than one POS were excluded from the list because their Yes/No test results would be difficult to interpret. For example, the adjective *abstract* had been selected from the AVL in our lemma sample; this item was excluded from the English cognates shortlist because there is also the noun *abstract*. 36 lemmas were excluded from the list after this POS check.

The 64 remaining words were checked for cognateness in the Find-A-Cognate database[7] to identify candidate words for the English Yes/No test. The Find-A-Cognate database comprises around 20,000 Spanish-English cognates (Montelongo et al. 2011). 53 words were cognates according to this database. We checked whether any of them had more than one Spanish translation equivalent. Six lemmas had more than one Spanish cognate and were excluded from the list of candidate words. For example, *extensive* was excluded because it maps onto both *extenso* and *extensivo*.

An attempt was made to match for POS and SUBTLEX-UK word frequency (van Heuven et al. 2014) the 47 shortlisted candidate English cognate words with non-cognate words selected from the AVL. Because frequency-matched non-cognates could not be found for all 47 cognate words, this matching activity yielded a list of 30 cognate and 30 non-cognate candidate words.

[6] Lexico is available at https://www.lexico.com/en/.

[7] The Find-A-Cognate database can be accessed at www.angelfire.com/ill/monte/findacognate.html.

Four fluent speakers of Spanish and English were given this list of 60 English words, each in a different random order. They were all native speakers of Spanish; three are PhD students and one is a lecturer at a British university. They were first given this definition of *cognate* in the instructions: 'Two words are called cognate if they belong to two languages and have similar form (i.e., spelling and sound) and meaning, such as English *castle* and Spanish *castillo*.' Then they were asked to tick the 'Yes' box below a word if they thought this word has a Spanish cognate word and the 'No' box below a word if they thought this word does not have a Spanish cognate word. The words that we considered cognate candidates based on the Find-a-cognate database search results and which were categorised as cognates by at least three respondents were shortlisted for inclusion in the English Yes/No test; the non-cognate candidates categorised as non-cognates by at least three respondents were shortlisted for inclusion in the English Yes/No test. Based on questionnaire responses one cognate candidate (*vessel*) and three non-cognate candidates (*settle*, *barrier*, *scarce*) were excluded from the item shortlist, together with their frequency-matched non-cognate and cognate words, respectively.

Finally, since pseudowords should follow the phonotactic and spelling regularities of a language to function as possible distractors, the pseudowords included in the Spanish and English Yes/No tests were created through the Spanish and English version of the software programme Wuggy (Keuleers and Brysbaert 2010) respectively. Wuggy[8] generates pseudowords which follow the phonotactics of specific languages.

This test used verbatim the instructions of the Yes/No test used in Schmitt et al. (2011). These instructions asked participants to check a word if they thought they can 'understand it when reading' and leave the gap next to the word blank if they thought they did not know a word (Schmitt et al. 2011, Supporting information file). The instructions

[8] Wuggy can be downloaded from http://crr.ugent.be/programs-data/wuggy.

also warned participants that the test included many non-words and familiarised participants with the task through examples and a practice phase.

4.2.3 Spanish word Yes/No test

When examining the possibility of cognate facilitation, it is necessary to check whether participants are familiar not only with the L2 cognate words but also with their L1 equivalent words (Daulton 2008; Nagy et al. 1993). A Spanish Yes/No test checked participants' familiarity with the Spanish cognates corresponding to the cognate words in the English Yes/No test.

The Spanish Yes/No test comprised a) the 26 Spanish loanwords corresponding to the English cognate words in the English Yes/No test, b) 26 Spanish words with various frequencies according to SUBTLEX-ESP, a word form frequency list from a corpus of Spanish subtitles (Cuetos et al. 2011), and c) 35 nonwords created by the Spanish version of Wuggy (Keuleers and Brysbaert 2010).

The same test instructions as in the English word Yes/No test were used but with Spanish word example and practice items. These example and practice items appear in Appendix B.

4.2.4 Procedure

Participants did all vocabulary tests in writing except LexTALE, which was done on computers. They did the tests in groups of 5 to 10 people in computer labs at two universities in Spain under the first author's supervision. After reading the participant information form and completing the consent form, participants responded to the data collection instruments in this order: English Yes/No test, LexTALE, Spanish Yes/No test and language learning background questionnaire. The English Yes/No test was administered before the Spanish Yes/No test because otherwise participants might guess the meaning of the English cognate words by associating them with their equivalent words in the Spanish Yes/No test. There was no time limit for any test. Participation lasted between 15 and 25 minutes.

5 Results

The Spanish Yes/No test results indicated that all participants knew all the Spanish cognates. The data of four participants were excluded from further analysis because they checked more than four pseudowords as known words in the English Yes/No test.

Target word frequency was operationalised as British English subtitle word form frequency in SUBTLEX-UK (see section 4 for a justification of this methodological choice). Form counts per million tokens were the frequency measure adopted from the SUBTLEX-UK database. The log10(frequency per million tokens) [*log10(SUBTLEX-UK frequency)*] of each word was calculated to minimise random variance.

Table 1 presents descriptive statistics for all variables. Because word frequency per million words is more easily interpretable than its log10 transformation, descriptive statistics for both measures are presented.

As mentioned in section 4.2.2, cognate and non-cognate words in this test were matched for frequency. Cognates and non-cognates indeed did not differ in terms of frequency according to a t-test comparing their log10(SUBTLEX-UK frequency) (t(50)

= .11, p = .916). Cognate words were on average longer than non-cognate words according to all measures of word length. Mann-Whitney tests comparing word length measures between cognates and non-cognates indicate that this finding was statistically significant, U(number of letters) = 151, z = -3.45, p = .001; U(number of phonemes) = 81, z = -4.74, p <.001 ; U(number of syllables) = 100, z = -4.53, p <.001.

Table 1 about here

A stepwise multiple regression analysis was conducted to examine which combination of the predictor variables can predict most of the IF variance. The next two paragraphs will summarise the correlation analysis findings that led us to choose this kind of regression analysis.

The relationships among predictor variables and between each predictor variable and IF, both overall and for cognate and non-cognate words separately, were explored to inform the design of the multiple regression analysis. Because not all variables had normally distributed data, Spearman rho correlations were conducted. Table 2 presents the results of these correlation analyses.

Table 2 about here

Word length measures correlated significantly and highly with each other. Word frequency did not correlate significantly with any word length measure for both cognate and non-cognate words.

Word frequency correlated significantly with IF overall but this correlation was stronger for non-cognate than for cognate words. This finding suggests an interaction between cognateness and frequency, whereby if an English word does not have a Spanish cognate, frequency correlates with IF more than if it has a Spanish cognate. Therefore, a multiple regression analysis which included a 'cognateness by frequency' interaction term was conducted to examine whether the effect of frequency on IF is moderated by whether a word is a cognate or not. Frequency was centred to enable the interpretation of the frequency regression coefficient and to decrease the multicollinearity which unavoidably occurred due to the correlation between frequency and the interaction term which includes frequency (Cohen et al. 2003: 201).

Data were analysed in SPSS 23. Predictors were entered in the regression analysis using the stepwise method, whereby all predictors are entered in the analysis and the analysis yields the most parsimonious regression model. The stepwise method was chosen over others because it would lead to the exclusion of any of the intercorrelated predictors which do not significantly affect the outcome variable.

Tests were conducted to examine whether the assumptions of multiple regression analysis were met. First, data were scanned for outliers through centred leverage (to examine whether any data points were unusual in any of the predictor variables), Cook's D (to examine the distance of data points in the outcome variable from their predicted values) and casewise diagnostics (to identify any data points in the outcome variable which are higher or lower than three SDs from the mean score). Centred leverage's maximum score was .14; given the sample size, a score larger than .98 would indicate an outlier. Cook's D maximum score was .16, which was below 1, the cutoff point for outliers. Casewise diagnostics identified the non-cognate word *livestock* as an outlier because it received an IF score 3.02 SDs below the mean IF score. Therefore, the data for *livestock* and its matched-for-frequency cognate word, *interpretation*, were excluded from the stepwise multiple regression analysis.

Following the deletion of these two items from the dataset, a Kolmogorov-Smirnov test conducted on the residuals ($D(50) = .8$, $p = .2$) and the histogram of the standardised residuals indicated that the data contained approximately normally distributed errors. In terms of the multicollinearity assumption, as mentioned above, a stepwise regression was conducted so that highly intercorrelated predictors which also did not contribute significantly to the variance of the outcome variable would be dropped from the final model. In this final model (see Table 3) tolerance statistics for all predictor variables were higher than .2, thus indicating that the assumption of no multicollinearity was met (Centred log of SUBTLEX-UK frequency, Tolerance = .51; Cognateness, Tolerance = 1; Centred log of SUBTLEX-UK frequency by Cognateness, Tolerance = .51). The conclusion that the assumption of no multicollinearity was met was also supported by the loading of predictors on the smallest eigenvalue because most of the variance of only the cognateness variable was related to this eigenvalue (Centred log of SUBTLEX-UK frequency, variance proportion: .30; Cognateness, variance proportion: .17; Centred log of SUBTLEX-UK frequency by Cognateness, variance proportion: .29). The assumption of independent errors was also met (Durbin-Watson = 1.58). The scatterplot between the studentised residuals and the predicted standardised residuals indicates randomly scattered data points without any curvature or funnel shape; therefore, this scatterplot suggests that the assumptions of homoscedasticity and linearity were met. The data were also examined for linearity via partial plots, which showed no sign of non-linear relationship between any of the predictor variables and the outcome variable.

Instead of following a rule of thumb to estimate the minimum sample size required in our analysis, we conducted a power analysis because rules of thumb 'do not take into account issues such as the expected effect size or the desired power of the test.' (Miles and Shevlin 2001: 119). In other words, it is impossible to assess how appropriate a rule of thumb is to a specific multiple regression analysis. An R^2 of .496 was our expected effect size because Willis and Ohashi's (2012) multiple regression analysis with similar predictor variables and outcome variable yielded this R^2. We used Cohen et al.'s (2003: 93) power analysis formula and assumed an α level of .05 and, as recommended by Cohen (1988), 80% power. We also specified six predictors in the formula since the regression model would have frequency, cognateness, number of letters, number of phonemes, number of syllables, and the interaction term between frequency and cognateness as predictors. According to the power analysis, 21 participants would be required. Therefore, the sample size in the present study (N = 34) was considered more than sufficient for this model.

Table 3 summarises the best regression model according to the stepwise multiple regression analysis. This model accounts for 44% of the IF variance ($R^2 = .44$, $R^2_{Adjusted} = .4$). The R^2 change between this model and the one without the interaction term (ΔR^2 = .07) is significant (F (1, 46) = 5.79, p = .02).

Table 3 about here

According to this model, frequency is the most important predictor, followed by cognateness and the interaction of frequency and cognateness. The stepwise regression had dropped word length measures from this model.

Coefficients have a different interpretation in a regression with an interaction term than in a regression without one (Cohen et al. 2003: 377). Because regressions with interaction terms are not often reported in applied linguistics articles, the interpretation of coefficients in this regression analysis will be reported. The coefficient for frequency indicates a predicted increase of .12 in IF when frequency increases by 1 among non-cognate words with average frequency. The coefficient for cognateness indicates that for words with average frequency, cognate words are predicted to have IF .06 higher than non-cognates. The coefficient of the interaction term suggests that frequency is predicted to lead to .08 less IF for a cognate than for a non-cognate.

The second research question examines the importance of each predictor in the multiple regression analysis. Squared semipartial correlations were calculated to address this research question. Table 4 presents them and their total.

Table 4 about here

Table 4 indicates that frequency uniquely predicts most of the variance in the outcome variable. Each of the other predictors makes a small unique contribution.

The squared semipartial correlations add up to .48, that is, .04 more than the R^2 of the best multiple regression model. In general, possible reasons for such findings are multicollinearity or suppression (Cohen et al. 2003: 425) but tolerance and eigenvalue statistics for the best regression model (summarised earlier in this section) indicated lack of multicollinearity. Therefore, we examined whether this finding is an indication of the existence of one or more *suppressor variables*. A suppressor variable is ‘a variable which increases the predictive validity of another variable (or set of variables) by its inclusion in the regression equation.' (Conger 1974: 36-37). In particular, we examined whether this finding is an indication of the suppression Conger (1974) called *reciprocal* and Cohen and Cohen (1975) called *cooperative*. In this kind of suppression a) two predictors are negatively correlated to each other and each is positively correlated to the outcome variable (or vice versa) and b) each predictor has a higher standardised regression coefficient when the other predictor is included in the regression analysis than when it is not. In the present study, frequency and cognateness were considered as likely reciprocal suppressors because they meet criterion a) of reciprocal suppression: they correlate negatively [r_s(Centred log10(SUBTLEX-UK frequency), cognateneness)= -.02, *p*= .92] and each of them correlates positively with IF [r_s(IF, Centred log10(SUBTLEX − UK frequency) = .046, *p* = .001; r_s(IF, cognateness)= .27, *p* = .056][9]. To test whether criterion b) is met we followed the procedure suggested by Paulhus et al. (2004). A single regression was conducted with each predictor and IF first, followed by a multiple regression where the other predictor was also entered in the analysis. We checked whether the standardised coefficient of the first predictor in each regression-analysis series increased after the entry of the second predictor. The standardised coefficient of centred log10(SUBTLEX-UK frequency) increased from .505 when it was the only predictor to .51 when the cognateness term had also been entered in the regression (ΔR^2 = .11). The standardised coefficient of cognateness increased from .32 when it was the only predictor to .33 when centred log10(SUBTLEX-UK frequency) had also been entered in the regression (ΔR^2 = .26). Consequently, frequency and cognateness are reciprocal suppressors. Because they are negatively correlated to each other and each of them is positively correlated to IF, when they both participate in the regression analysis they

[9] All these correlations indicate that criterion a) is met because it is not a requirement of this criterion that any of these correlations be statistically significant.

enhance each other's contribution to IF variance by suppressing IF-irrelevant variance in each other.

6 Discussion

This study examined how much three lexical characteristics can predict Spanish L1 university students' recognition knowledge of English academic vocabulary. Since this investigation was partly motivated by the contrasting findings of Milton and Daller (2007) and Willis and Ohashi (2012), the findings of the present study will be compared and contrasted with the findings of these studies.

In both this study and Milton and Daller (2007) target word frequency is the most important predictor of IF. However, unlike Milton and Daller (2007), the present study suggests that cognateness affects EFL academic vocabulary learning. Milton (2009: 41) attributed the lack of an effect for cognateness in Milton and Daller (2007) to the high percentage of cognate words between English and French or to their similar morphology, the rationale being that perhaps these factors diminished cognate facilitation. However, this study indicates that cognate facilitation can be statistically significant in language pairs rich in cognates and morphologically similar.

The findings of the present study also partly agree with those in Willis and Ohashi (2012). Both studies found a significant effect for cognateness and frequency but in Willis and Ohashi (2012) cognateness was the most important predictor and frequency the second most important predictor, whereas in this study their roles were reversed. The different cognateness and frequency effect sizes between the two studies can be due to one or more methodological differences between them. In Willis and Ohashi (2012) the number of cognates was not stable across the BNC frequency bands sampled in the VST, ranging from two in the third and sixth band to four in the first band, whereas in the present study cognates and non-cognates were matched for frequency. The lack of an equal number of cognates sampled across frequency bands in Willis and Ohashi (2012) may have confounded the roles of cognateness and frequency in the regression analysis. Another possible reason is the difference in participants' English proficiency level; based on vocabulary test scores – VST in Willis and Ohashi (2012) and LexTALE here – most of Willis and Ohashi's participants (2012) had a low-intermediate level whereas in our study most participants were of upper-intermediate and advanced level. Since the lower one's EFL proficiency level, the higher the cognate facilitation effect in receptive vocabulary tests (Laufer and McLean 2016), the higher contribution of cognateness to word learning in Willis and Ohashi (2012) than in our study may be due to their participants' lower proficiency level.

Contrary to Willis and Ohashi (2012), this study suggests a significant interaction between frequency and cognateness. The lack of a significant interaction in Willis and Ohashi (2012) may be due to the possible confounding between frequency and cognateness.

Our results are similar to Elgort's (2013), where the interaction between frequency and cognateness approached significance. We consider cognateness as the moderator in the interaction term in our regression because the effect of frequency on IF was larger for non-cognates than for cognates. By contrast, Elgort (2013: 269) considered frequency as the moderator because 'the difference between accuracy of responses to cognates and non-cognates decreased as item frequency increased'. Contrary to Elgort's comment, Figure 4 in Elgort (2013: 267) (see Appendix D) indicates that the difference in the accuracy of responses to cognates and non-cognates does not decrease as item frequency increases but only when the logarithm of frequency is between 2 and 3. Therefore, Figure 4 indicates that the relationship between frequency and accuracy is different for cognates and non-cognates. The curve for cognates is an inverse polynomial (that is, the accuracy scores increase steadily until a certain frequency score is reached, after which the growth rate diminishes) whereas for non-cognates it is cubic (that is, it changes direction at two points). Since this graph indicates that the relationship between frequency and accuracy differs between cognates and non-cognates, in Elgort (2013) cognateness moderates the relationship between frequency and accuracy.

Squared semipartial correlations suggested that 30% of the variance in the IF scores can be attributed to word frequency and that cognateness and the interaction between frequency and cognateness account for 11% and 7% of the IF variance, respectively. Because Willis and Ohashi (2012) calculated the unique contribution of their predictor variables not through squared semipartial correlations but by averaging the simple (otherwise called 'zero-order') correlations across the regression models created in the stepwise regression analysis they conducted, their results are not comparable to ours. Given the possible confounding between cognateness and frequency in their study, and the greater validity of squared semipartial correlations over simple correlations as a measure of the unique contribution of each predictor on an outcome variable (Tabachnick and Fidell 2014: 180), our findings are likely to be more informative than those of Willis and Ohashi (2012).

7 Limitations of the study

The lack of an effect of word length on IF is surprising given the significant negative relationship between word length and IF found in Willis and Ohashi (2012). This lack of a relationship between word length and IF in the present study may be due to the similar morphological makeup of English and Spanish; although affixes may render words ‘long’, thanks to regular interlingual equivalences, knowledge of these equivalences cancels out the effect of word length on EFL word learning. For example, the English adverbial suffix -*ly* corresponds to Spanish *–mente,* as in *considerably – considerablemente* (see Green and Coxhead 2015: 61-62 for more examples). Alternatively, this lack of a word-length effect on IF could be due to the fact that in this study cognate words were significantly longer than non-cognate words. In other words, because the mean vocabulary length is different for cognate and non-cognate words, cognateness and word length are, to some extent, confounded. Consequently, some of the IF variance actually due to word length may have been attributed wrongly to cognateness. Matching cognates and non-cognates not only for frequency but also for word length was impossible because we used real academic words to enhance the ecological validity of this study. However, future research with pseudocognates (see De Groot and Keijzer 2000) could ensure that cognate and non-cognate test items are matched for word frequency, cognateness and word length and, therefore, could examine thoroughly the relative effect of all these variables on IF.

The generalisability of our findings is limited to the most frequent 1,000 AVL lemmas because only words from the most frequent 1,000 AVL lemmas were tested. However, as mentioned in sections 4 and 4.2.2, sampling words only from the first AVL frequency band had its methodological benefits: it led to a sampling rate high enough to warrant the claim that the findings of our study are generalisable to the whole band while keeping the duration of the study manageable for participants. Besides, limiting the sampling to words inside only one frequency band did not mean that only very frequent words were tested; the frequency range of words in the English Yes/No test is wide, spanning from 0.95 to 80.47 occurrences per million words in SUBLTEX-UK (SD = 20.37) (see Table 1 for more details). Moreover, we assessed how the most frequent 1,000 AVL lemmas overlap with words in English reference corpora in terms of frequency by searching them in the 1-25 1,000 BNC-COCA word family bands[10]. This search was done via the VocabProfile software in the Compleat Lexical Tutor website[11]. 44.5% of the lemmas fall inside the first two BNC-COCA bands. The rest of the lemmas fall inside the third (44.3%), fourth (9%), fifth (1.5%) and seventh (0.3%) BNC-COCA bands. Four AVL lemmas do not appear in any of the 25 BNC-COCA bands. Therefore, the first AVL band mainly contains *high-frequency vocabulary* when this is defined as the 3,000 (not 2,000) most frequent word families (Schmitt and Schmitt 2014). It also contains what Schmitt and Schmitt (2014) have called *mid-frequency vocabulary*, that is, vocabulary between the fourth and ninth frequency bands.

The generalisability of our findings is also limited to academic vocabulary *form recognition* knowledge because a Yes/No test was used. It should also be pointed out that Yes/No tests have been criticised because, as all self-report measures, they are prone to overestimation or underestimation of participants' vocabulary knowledge. As in any study employing a Yes/No test, in this study it was impossible to limit the effect of possible knowledge underestimation by participants but we guarded against its overestimation through the inclusion of pseudowords and the exclusion from further statistical analyses of the answers of participants who ticked more than 4 pseudowords as

---

[10] The procedure followed to compile these bands is summarised at https://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Information-on-the-BNC_COCA-word-family-lists.pdf.

[11] The VocabProfile software can be accessed at https://www.lextutor.ca/cgi-bin/vp/comp/output.pl.

known. We preferred to exclude participant data from further analysis instead of using one of the available penalisation formulae because it is unclear how effective they are (Beeckmans et al. 2001; Huibregtse et al. 2002).

The sample size of our study may be considered small according to some of the rules of thumb on vocabulary size in multiple regression analysis, such as the one suggesting 10 participants per predictor variable (Howell 2002). However, many 'statisticians are sceptical about any rules of thumb' for sample size estimation (Levshina 2015: 144) and recommend conducting power analysis instead because they yield estimates that are tailor-made to the expected effect size and assumed power of each specific multiple regression analysis (e.g., Miles and Shevlin 2001). The sample size of this study is more than sufficient because it is higher than the optimal sample size indicated by power analysis.

10 participants had spent at least three months in an English-speaking country (see section 4.1), a period of time that could have led to a considerable increase in the number of English words they could recognise. Because we do not know whether the proportion of participants who had spent a considerable amount of time in an English-speaking country is generalisable to the whole population of university students in Spain, we estimated the possible impact that the inclusion of these participants had to the regression analysis findings. To this end, we conducted a stepwise regression analysis with the same predictors and outcome variable as the one conducted to address Research Question 1 but only with the data of the 26 participants who had both spent less than three months in an English-speaking country and had not checked more than four pseudowords in the English Yes/No test[12]. Casewise diagnostics indicated that the non-cognate words

[12] This regression analysis was conducted with the data of eight rather than 10 participants less than those in the main regression analysis because the data of two of the participants who had spent more than three

livestock and *stance* were outliers because the former received an IF score 3.06 SDs below the mean IF score and the latter received an IF score 3.04 SDs below the mean IF. Therefore, the final regression analysis was conducted after excluding the data for these words and their frequency-matched non-cognate words (*interpretation*, *discrimination*). All multiple regression assumptions were met. The assumption test results are summarised in Appendix E.

The final model yielded by the stepwise analysis accounts for 49.8% of the IF variance ($R^2 = .498$, $R^2_{Adjusted} = .463$), an effect size very cloze to that of the regression analysis with 34 participants' data ($R^2 = .44$, $R^2_{Adjusted} = .4$). Moreover, according to both the present analysis and the one with 34 participants' data, the same predictors (frequency, cognateness and their interaction) had a significant effect on IF. The coefficients of the predictors in both analyses were very similar, as a comparison between Table 3 and Table 5 indicates. The unique contribution of each predictor to IF was also very similar between the two analysis, as a comparison between Table 4 and Table 6 indicates. Consequently, the results of the main regression analysis reported in this study (see section 5) had not been distorted by the data of participants who had spent three months or more in an English-speaking country.

Table 5 about here

months in an English-speaking country had already been excluded from the original regression analysis (see section 5) because they had checked more than four pseudowords as known words in the English Yes/No test.

Table 6 about here

8 Pedagogical implications

Our finding that cognateness can offer an advantage even to advanced-level EFL Spanish L1 learners indicates that helping Spanish L1 learners identify cognates is worthwhile when considered in combination with these findings:

a) Spanish-English bilinguals do not always recognise cognates as such (e.g., Nagy et al. 1993; August et al. 2005), not all Spanish-English bilingual children learn cognates more easily than non-cognates (e.g., Kelley and Kohnert 2012; Potapova et al. 2016),
b) even adult upper-intermediate and advanced-level EFL Spanish L1 learners do not recognise all English academic-word cognates since only 10 out of the 26 cognates in the English Yes/No test were recognised by all participants in our study and
c) awareness-raising activities about cognates are effective (Nagy et al. 1993; Proctor and Mo 2009).

Awareness-raising activities about cognates are recommended for words with L1-L2 equivalent affixes and roots. Green and Coxhead (2015) and Garrison (1990) suggest presenting learners with lists of Spanish-English equivalent suffixes (e.g., *-ista –ist*, as in *novelista*, *novelist*; *-oso –ous*, as in *famoso*, *famous*). Learning about these equivalent suffixes will enable Spanish-English bilinguals to recognise cognates during reading and listening. Garrison (1990: 511) also suggests a follow-up activity aiming to help learners develop confidence with word derivation; he suggests giving pupils the root of an L1 cognate word and asking pupils to guess the equivalent L2 word by using the right L2 suffix. Hernández and Montelongo (2018) suggest another follow-up activity where learners are asked to guess at the meaning of words which end in a suffix that is similar between English and Spanish (e.g., *generous*). They also suggest word sorting activities to raise learners' awareness not only of suffixes but also prefixes (e.g., *uni-* as in *unicornio, unicorn*; *pre-* as in *prediction, predicción*) and roots (e.g., *graph* (meaning *write*) and *auto* (meaning *self*) as in *autograph, autografia*) of Latin and Greek origin. They recommend this sequence of sorting activities: first students are given a list of English words and a list of their Spanish cognates and they match the cognates, then they group the English-Spanish pairs per suffix/prefix/root, and then they discuss in class about the meaning of the suffix/prefix/root and the meaning of the cognate words.

Cognate-awareness raising activities also lead to better word-meaning inferencing and, consequently, reading comprehension (Proctor and Mo 2009). Such activities seem useful because research suggests that not all Spanish-English bilingual primary-school children strategically use their knowledge of cognates during reading in English (e.g., Jiménez et al. 1996; García 1998). Teaching intervention research involving cognate-awareness raising activities suggests that such activities combined with reading comprehension strategy instruction can help Spanish-English bilingual children recognise cognates during reading and infer their meaning (e.g., Dressler et al. 2011).

Cognate awareness raising activities are impossible when there is little spelling and/or phonological overlap between cognate word pairs. For example, in our pilot study the bilingual judges whom we consulted in our study did not think that *vessel* is a cognate of a Spanish word although according to the Find-a-cognate database *vessel* is etymologically related to *vasija*. Academic words which differ considerably from their etymological cognates may require direct instruction (Lubliner and Hiebert 2011: 88), especially if they are also infrequent and, therefore, unlikely to be encountered enough times by EFL learners so as to be learned incidentally. Rodriguez (2010: 565) points out that cognates which are more similar phonetically than orthographically may be particularly difficult to recognise (e.g., *peace*, *paz*). He suggests introducing such cognates in listening activities where pupils are asked to infer their meaning from context.

Viewed from a language testing perspective, this study suggests that at least when upper-intermediate Spanish EFL learners' receptive knowledge of the most frequent 1,000 AVL lemmas is tested, frequency has a far greater effect on scores than cognateness and, consequently, cognateness is unlikely to inflate vocabulary size scores. However, the greater effect of cognates on test scores for L1 Japanese than for L1 Russian EFL learners in Laufer and McLean (2016) suggests that the influence of cognates on test scores may vary across L1s; examining the reasons behind these L1 effects is necessary before drawing conclusions on whether cognates should be included in vocabulary tests or not.

9 Conclusion and implications for research

The present study provides further evidence about the importance of word frequency in L2 vocabulary learning (e.g., Elgort 2013) and the advantage that cognates have over non-cognates (e.g., Laufer and McLean 2016; Willis and Ohashi 2012). Both findings are important because, unlike Willis and Ohashi (2012) and Petrescu et al. (2017), this study measured the *separate* effect of frequency and cognateness on word learnability thanks to its approach to word sampling for the English Yes/No test (see section 4.2.2).

Apart from corroborating earlier research findings, the present study yielded novel findings. The statistically significant interaction between frequency and cognateness suggests that the effect of frequency on IF is greater for non-cognates than cognates. Since in Elgort (2013) a frequency by cognateness interaction only approached significance, similar studies are needed to examine whether this finding is replicable. The reciprocal suppression between word frequency and cognateness suggests that these word characteristics boost each other's contribution to word learnability. Future research should examine whether this suppression is replicable.

This study's findings are generalisable to the development of advanced and upper-intermediate level Spanish L1 EFL learners' ability to recognise academic vocabulary learning, at least as far as Yes/No tests can indicate one's ability to recognise vocabulary and LexTALE scores can indicate one's EFL proficiency level. The kind of vocabulary test administered and learners' proficiency level modulate cognate facilitation because cognate facilitation increases the more difficult a test is (see Laufer and Goldstein 2004) and the lower the participants' proficiency level (Laufer and McLean, 2016). Therefore, studies with Spanish L1 EFL learners of other proficiency levels and using not only recognition tests are necessary.

The present study tested words sampled from the most frequent 1,000 lemmas of the AVL (Gardner and Davies 2014). In Petrescu et al. (2017) cognate facilitation increased the rarer the words; it would be interesting to see whether testing Spanish L1 university students on a sample from all AVL lemmas would replicate Petrescu et al.'s (2017) finding. A more complicated finding is suggested by Bennet and Stoeckel (2014a), where infrequent L1 loanwords were not a source of DIF and by Potapova et al. (2016: 724), where a cognate advantage was not found for infrequent L1 loanwords in Spanish dominant Spanish-English adult bilinguals' performance in the *Peabody Picture Vocabulary Test-Third Edition (PPVT-III)* (Dunn and Dunn 1997). A study testing a sample of all AVL lemmas could suggest an increase in cognate facilitation the rarer the lemmas are but once frequency becomes so low that participants do not know the L1 cognates of the target words, cognate facilitation could decrease. Moreover, since Bennett and Stoeckel (2014a; 2014b) claim that loanword frequency affects cognate recognition and various studies have indicated that many Spanish words are more frequent than their English cognates (e.g., Lubliner and Hiebert 2011; Chen et al. 2012), examining the extent to which L1 Spanish word frequency affects English cognate word learning and vice versa is another avenue for future research.

References

Allen, David B. & Kathy Conklin. 2013. Cross-linguistic similarity and task demands in Japanese-English bilingual processing. *PLOS ONE* 8(8). 1–14.

Alsaif, Abdullah & James Milton. 2012. Vocabulary input from school textbooks as a potential contributor to the small vocabulary uptake gained by English as a foreign language learners in Saudi Arabia. *The Language Learning Journal* 40(1). 21–33.

August, Diane, Maria Carlo, Cheryl Dressler & Catherine Snow. 2005. The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice* 20(1). 50–57.

Beeckmans, Renaud, June Eyckmans, Vera Janssens, Michel Dufranne & Hans Van de Velde. 2001. Examining the Yes/No vocabulary test: some methodological issues in theory and practice. *Language Testing* 18(3). 235–274.

Beglar, David. 2010. A Rasch-based validation of the vocabulary size test. *Language Testing* 27(1). 101–118.

Bennett, Phil & Tim Stoekel. 2014a. Does loanword frequency predict DIF in L2 vocabulary tests? Poster presented at the Japan Association for Language Teaching annual PanSIG Conference, Miyazaki.

Bennett, Phil & Tim Stoeckel. 2014b. Word frequency and frequency of loanwords as predictors of word difficulty. *VERB* 3(2). 4–5.

Brown, James D. 2005. *Testing in language programs: A comprehensive guide to English language assessment* (2nd edn.). New York: McGraw-Hill.

Brysbaert, Mark & 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(4). 977–990.

Chen, Zheng & Grant Henning. 1985. Linguistic and cultural bias in language proficiency tests. *Language Testing* 2(2). 155–163.

Chen, Xi, Gloria Ramirez, Yang C. Luo, Esther Geva & Yu-Min Ku. 2012. Comparing vocabulary development in Spanish and Chinese-speaking ELLs: the effects of metalinguistic and sociocultural factors. *Reading and Writing* 25(8). 1991–2020.

Cobb, Tom. 2010. Learning about language and learners from computer programs. *Reading in a Foreign Language* 22(1). 181–200.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ Lawrence Erlbaum Associates, Publishers.

Cohen, Jacob & Patricia Cohen. 1975. *Applied multiple regression/correlation analysis for the behavioural sciences.* New York: Routledge.

Cohen, Jacob, Patricia Cohen, Stephen G. West & Leona S. Aiken. 2003. *Applied multiple regression/correlation analysis for the behavioural sciences.* New York: Routledge.

College Entrance Examination Center. 2002. Reference word list for senior high English education. http://www.ceec.edu.tw/Research2/Default.aspx. (accessed March 2019)

Coxhead, Averil. 2000. A new Academic Word List. *TESOL Quarterly* 34(2). 213–238.

Csomay, Eniko & Alexandra Prades. 2018. Academic vocabulary in ESL student papers: A corpus-based study. *Journal of English for Academic Purposes* 33. 100–118.

Cuetos, Fernando, Maria Glez-Nosti, Analía Barbón & Marc Brysbaert. 2011. SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica* 32. 133–143.

Culligan, Brent. 2015. A comparison of three test formats to assess word difficulty. *Language Testing* 32(4). 503–520.

Daulton, Frank E. 2005. Common *Gairaigo* loanwords corresponding to high-frequency and academic vocabulary: Are our students ready for foreign study? *JALT Hokkaido Journal* 9. 1–16.

Daulton, Frank E. 2008. *Japan's built-in lexicon of English-based loanwords*. Clevedon. London: Multilingual Matters.

De Groot, Annettte M.B. & Rineke Keijzer. 2000. What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign language learning and forgetting. *Language Learning* 50(1). 1−56.

Demetriou, Lizeta. 2017. Selection, Presentation and Recycling of Phrasal Verbs in ESL Textbooks in Cyprus. Paper presented at the 2017 BAAL Vocabulary Special Interest Group conference, Reading, UK.

Dressler, Cheryl, Maria S. Carlo, Catherine E. Snow, Diane August & Claire E. White. 2011. Spanish-speaking students’ use of cognate knowledge to infer the meaning of English words. *Bilingualism: Language and Cognition* 14(2). 243–255.

Dunn, Lloyd M. & Douglas M. Dunn. 1997. *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Durrant, Philip. 2016. To what extent is the Academic Vocabulary List relevant to university student writing? *English for Specific Purposes* 43. 49–61.

Ecke, Peter & Christopher J. Hall. 2014. The parasitic model of L2 and L3 vocabulary acquisition: evidence from naturalistic and experimental studies. *Fórum Linguístico* 11(3). 360–372.

Elgort, Irina. 2013. Effects of L1 definitions and cognate status of test items on the Vocabulary Size Test. *Language Testing* 30(2). 253–272.

Eyckmans, June. 2004. *Measuring receptive vocabulary size*. Utrecht: LOT.

Fellbaum, C. (ed.) 1998. *WordNet: An Electronic Lexical Database.* Cambridge, MA: MIT Press.

García, Georgia E. 1998. Mexican-American bilingual students' metacognitive reading strategies: What's transferred, unique, problematic? *National Reading Conference Yearbook 47*. 253–263.

Gardner, Dee. & Marc Davies. 2014. A new Academic Vocabulary List. *Applied Linguistics* 35(3). 305–327.

Greene, Jennifer & Averil Coxhead. 2015. *Academic vocabulary for middle school students. Research-based lists and strategies for key content areas.* Baltimore: Brookes Publishing.

Gyllstad, Henrik, Laura Vilkaité & Norbert Schmitt. 2015. Assessing vocabulary size through multiple-choice formats: Issues with guessing and sampling rates. *ITL-International Journal of Applied Linguistics* 166(2). 276–303.

Hernández, Anita C. & José A. Montelongo. 2018. Word study with English Spanish cognates. In Kathy Ganske (ed.) *Word sorts and more: Sound, pattern and meaning explorations K-3* (2nd edn.), 81–92. New York: Guilford Press.

Horst, Marlise, Tom Cobb & Paul Meara. 1998. Beyond A Clockwork Orange: Acquiring Second Language Vocabulary through Reading. *Reading in a Foreign Language* 11(2). 207–223.

Howell, David C. 2002. *Statistical methods for psychology*. Pacific Grove, CA: Duxbury/Thompson Learning.

Huibregtse, Ineke, Wilfried Admiraal & Paul Meara, P. 2002. Scores on a yes-no vocabulary test: Correction for guessing and response style. *Language Testing* 19(3). 227–245.

Hyland, Ken & Polly Tse. 2007. Is there an 'academic vocabulary'? *TESOL Quarterly* 41(2). 235–253.

Jiang, Nan. 2000. Lexical representation and development in a second language. *Applied Linguistics* 21(1). 47–77.

Jiang, Nan. 2004. Semantic transfer and its implications for vocabulary teaching in a second language. *The Modern Language Journal* 88(3). 416–432.

Jiménez, Robert T., Georgia E. García & P. David Pearson. 1996. The reading strategies of bilingual Latina/o students who are successful English readers: Opportunities and obstacles. *Reading Research Quarterly* 31(1). 90–112.

Kellerman, Eric. 1978. Giving Learners a Break: Native Language Intuitions as a Source of Predictions about Transferability. *Working Papers on Bilingualism* 15. 59–92.

Kelley, Alaina & Kathryn Kohnert. 2012. Is there a cognate advantage for typically developing Spanish-speaking English-language learners? *Language, speech, and hearing services in schools* 43(2). 191–204.

Keuleers, Emmanuel & Marc Brysbaert. 2010. Wuggy: A multilingual pseudoword generator. *Behavior Research Methods* 42(3). 627–633.

Kieffer, Michael J. & Catherine DiFelice Box 2013. Derivational morphological awareness, academic vocabulary, and reading comprehension in linguistically diverse sixth graders. *Learning and individual differences* 24. 168–175.

Laufer, Batia. 1990. Why are Some Words More Difficult than Others? *IRAL: International Review of Applied Linguistics in Language Teaching* 28(4). 293–307.

Laufer, Batia. 1997. What's in a word that makes it hard or easy: Some intralexical factors that affect the learning of words. In Norbert Schmitt & Michael McCarthy (eds.), *Vocabulary: Description, acquisition and pedagogy*, 140–155. Cambridge: Cambridge University Press.

Laufer, Batia & Zahava Goldstein. 2004. Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning* 54(3). 399–436.

Laufer, Batia & Stuart McLean. 2016. Loanwords and Vocabulary Size Test Scores: A Case of Different Estimates for Different L1 Learners. *Language Assessment Quarterly* 13(3). 202–217.

Lemhöfer, Kristin & Mirjam Broersma. 2012. Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods* 44(2). 325–343.

Levshina, Natalia. 2015. *How to do statistics with R: Data exploration and statistical analysis.* Amsterdam: John Benjamins.

Lubliner, Shira & Elfrieda H. Hiebert. 2011. An analysis of English–Spanish cognates as a source of general academic language. *Bilingual Research Journal* 34(1). 76–93.

Macis, Marijana and Norbert Schmitt. 2017. Not just 'small potatoes': Knowledge of the idiomatic meanings of collocations. *Language Teaching Research* 21(3). 321–

340.

Masrai, Ahmed & James Milton. 2018. Measuring the contribution of academic and general vocabulary knowledge to learners' academic achievement. *Journal of English for Academic Purposes* 31. 44–57.

Meara, Paul. 1992. *EFL Vocabulary Tests.* University College Swansea: Centre for Applied Linguistics Studies.

Meara, Paul & Barbara Buxton. 1987. An alternative to multiple choice vocabulary tests. *Language Testing* 4(2). 142–154.

Melitz, Jacques. 2018. English as a lingua franca: Facts, benefits and costs. *The World Economy* 41.1750–1774.

Miles, Jeremy N. V. and Mark Shevlin. 2001. *Applying regression and correlation: A guide for students and researchers.* London: Sage.

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

Milton, James & Helmut Daller. 2007. The interface between theory and learning in vocabulary acquisition. Paper presented at EUROSLA 2007, Newcastle upon Tyne, UK.

Montelongo, José. A., Anita Hernández & Roberta Joan Herter. 2011. Identifying Spanish–English cognates to scaffold instruction for Latino ELs. *The Reading Teacher* 65(2). 161–164.

Nagy, William E, Georgia Earnest García, Aydin Y. Durgunogiu & Barbara Hancin-Bhatt. 1993. English–Spanish bilingual students' use of cognates in English reading. *Journal of Reading Behavior* 25(3). 241– 259.

Nash, Rose. 1997. *NTC's dictionary of Spanish cognates.* Chicago, IL: NTC Publishing Group.

Nation, I. S. Paul. 1990. *Teaching and learning vocabulary.* Florence, KY: Heinle and Heinle.

Nation, I. S. Paul. 2012. The BNC/COCA word family lists (17 September 2012). Unpublished paper. http://www.victoria.ac.nz/lals/about/staff/paulnation (accessed June 2016).

Nation, I. S. Paul & David Beglar. 2007. A vocabulary size test. *The Language Teacher* 31(7). 9–13.

Otwinowska, Agnieszka & Jakub M. Szewczyk. 2017. The more similar the better? Factors in learning cognates, false cognates and non-cognate words. *International Journal of Bilingual Education and Bilingualism*. DOI: 10.1080/13670050.2017.1325834.

Paquot, Magali. 2010. *Academic vocabulary in learner writing. From extraction to analysis.* London: Continuum.

Paribakht T. Sima & Marjorie Wesche. 1997. Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In James Coady & Thomas Huckin (eds.), *Second language vocabulary acquisition,* 174–200. Cambridge: Cambridge University Press.

Pellicer-Sánchez, Ana. 2017. Learning L2 collocations incidentally from reading. *Language Teaching Research* 21(3). 381–402.

Peters, Elke & Stuart Webb. 2018. Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition* 40. 551–577.

Petrescu, Maria Claudia, Rena Helms-Park & Vedran Dronjic. 2017. The impact of frequency and register on cognate facilitation: Comparing Romanian and Vietnamese speakers on the Vocabulary Levels Test. *English for Specific Purposes* 47. 15–25.

Potapova, Irina, Henrike K. Blumenfeld & Sonja Pruitt-Lord. 2016. Cognate identification methods: Impacts on the cognate advantage in adult and child Spanish-English bilinguals. *International Journal of Bilingualism* 20(6). 714–731.

Proctor, C. Patrick & Elain Mo. 2009. The Relationship Between Cognate Awareness and English Comprehension Among Spanish–English Bilingual Fourth Grade Students. *TESOL Quarterly* 43(1). 126–136.

Reynolds, Barry Lee & David Wible. 2014. Frequency in incidental vocabulary acquisition research: An undefined concept and some consequences. *TESOL Quarterly* 48(4). 843–861.

Reynolds, Barry Lee, Wei-Hua Wu, Hui-Wen Liu, Shu-Yuan Kuo & Ching-Hua Yeh. 2015. Towards a model of advanced learners' vocabulary acquisition: an investigation of L2 vocabulary acquisition and retention by Taiwanese English majors. *Applied Linguistics Review* 6(1). 121–144.

Reynolds, Barry Lee, Ying-Chun Shih & Wei-Hua Wu. 2018. Modeling Taiwanese adolescent learners' English vocabulary acquisition and retention: The washback effect of the College Entrance Examination Center's reference word list. *English for Specific Purposes* 52. 47–59.

Sasaki, Miyuki. 1991. A comparison of two methods for detecting differential item functioning in an ESL placement test. *Language Testing* 8(2). 95–111.

Schmitt, Norbert. 2000. *Vocabulary in language teaching*. Cambridge: Cambridge University Press.

Schmitt, Norbert. 1998. Tracking the incremental acquisition of second language vocabulary: a longitudinal study. *Language learning* 48(2). 281–317.

Schmitt, Norbert, Xiangying Jiang & William Grabe. 2011. The percentage of words known in a text and reading comprehension. *The Modern Language Journal* 95(1). 26–43.

Schmitt, Norbert & Diane Schmitt. 2014. A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching* 47(4). 484–503.

Schuth, Elisabeth, Judith Köhne & Sabine Weinert. 2017. The influence of academic vocabulary knowledge on school performance. *Learning and Instruction* 49. 157–165.

Spencer, Sarah, Judy Clegg, Hilary Lowe & Joy Stackhouse. 2017. Increasing adolescents' depth of understanding of cross-curriculum words: an intervention study. *International journal of language & communication disorders* 52(5). 652–668.

Stoeckel, Tim & Phil Bennett. 2013. Sources of differential item functioning between Korean and Japanese examinees on a second-language vocabulary test. *Vocabulary Learning and Instruction* 2(1). 47–54.

Szudarski, Paweł & Ronald Carter. 2016. The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics* 26(2). 245–265.

Tabachnick, Barbara G. & Linda S. Fidell. 2014. *Using multivariate statistics.* Boston: Pearson.

Townsend, Dianna, Alexis Fillippini, Penelope Collins and Gina Biancarosa 2012. Evidence for the Importance of Academic Word Knowledge for the Academic Achievement of Diverse Middle School Students. *The Elementary School Journal* 112(3). 497–518.

Truckenmiller, Adrea J. & Yaacov Petscher. 2019 (online first) The role of academic language in written composition in elementary and middle school. *Reading and Writing*. DOI: 10.1007/S11145-019-09938-7

van Heuven, Walter J.B., Pawel Mandera, Emmanuel Keuleers & Marc Brysbaert. 2014. SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology* 67(6). 1176–1190.

Vidal, Karina. 2003. Academic listening: a source of vocabulary acquisition? *Applied Linguistics* 24(1). 56–89.

West, Michael. 1953. *A General Service List of English Words.* London: Longman, Green & Co.

Willis, Martin & Yoshie Ohashi. 2012. A model of L2 vocabulary learning and retention. *The Language Learning Journal* 40. 125–137.

Zumbo, Bruno D. 1999. *A handbook on the theory and methods of differential item functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Type (Ordinal) Item Scores.* Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense. http://faculty.educ.ubc.ca/zumbo/DIF/ (accessed on 19 September 2019)

Appendix A: Items in the English Yes/No test

Cognate words	**SUBTLEX-UK frequency-matched non-cognate words**
university (S: universidad)	growth
require (S: requerir)	rely
response (S: respuesta)	knowledge
central (S: central)	available
apply (S: aplicar)	seek
association (S: asociación)	wealth
attitude (S: actitud)	degree
interaction (S: interacción)	likelihood

interpretation (S: interpretación)	livestock
evaluate (S: evaluar)	broaden
definition (S: definición)	workshop
percentage (S: porcentaje)	tool
typical (S: típico)	useful
discovery (S: descubrimiento)	belief
myth (S: mito)	ownership
discrimination (S: discriminación)	stance
restriction (S: restricción)	thinker
specify (S: especificar)	outweigh
infrastructure (S: infraestructura)	outcome
inclusion (S: inclusión)	drawback
hierarchy (S: jerarquía)	offspring
conversion (S: conversión)	allowance
flexibility (S: flexibilidad)	weakness
substantially (S: substancialmente)	lastly
considerably (S: considerablemente)	nonetheless
justification (S: justificación)	outset

Note: Spanish cognates appear within parentheses.

Pseudowords

haddy, nitch, dreas, cag, halm, dracer, cround, cround, bood, stad, jolder, sping, kile, totle, hode, craddock, sporly, verden, poot, cridge, plany, pernicate, treak, repow, witten, earch, enruy, skelding, gurl, jink, lannery, casning, sistence, thint, snurley.

Appendix B: Example and practice items for the Spanish word Yes/No test

Note: The feedback about the practice items appeared on a different page in the handout.

If you know this word ___√___ perro

If you do not know this word _____ imperceptible

If you do not know this word _____ mintar

(Good, because it is a non-word)

If you check a non-word, you will lose points __√__ mintar

Now try some practice words

_____ coche

_____ remojado

_____ tinfeta

_____ fealdad

_____ chismear

_____ erto

_____ día

_____ fasmoso

_____ conocer

_____ obsequiar

After completing the practice words above, you might have checked any of 7 real words that you know. But you should not have checked "tinfeta," "erto" o "fasmoso" because they are not real words in Spanish.

Appendix C: Language Background Questionnaire

In this short questionnaire you will be asked about your personal details and about your experience learning English as a foreign language. This test is comprised by 15 questions which do not require long answers. You must answer all the questions that are applicable to you. There are not right or wrong answers to these questions as each person may have a different language learning experience. Completing this questionnaire will not take you more than 3-4 minutes.

Please answer the following questions.

Personal details

1. Name and surname:

2. Age (in years): _____

3. Sex (circle one): Male / Female

4. Education (degree obtained or school level attended/ if you are currently studying specify what):

5. Country of origin:

6. Country of Residence:

7. If your answers to questions 5 and 6 are the same, have you travelled or lived abroad in a country where your second language (English) is spoken? Where? How long for?

8. If your answers to questions 5 and 6 are different, how long have you been in the country of your current residence (in years)?

Language background and experience

9. What is your *native* language, that is, the language you first spoke? If there is more than one, please list them.

10. Do you know any other languages in addition to your native language(s) and English?

11. If you answered 'Yes' to question 10, please list this language or these languages.

Appendix D: Figure 4 in Elgort (2013: 267), reproduced with permission

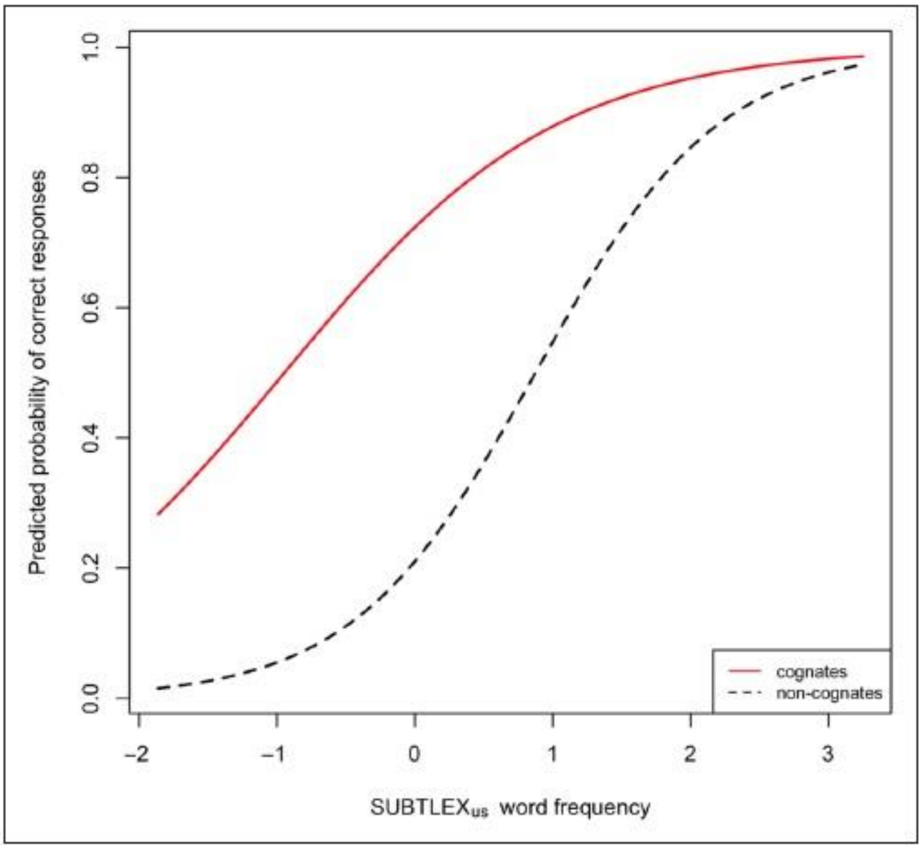

Figure 4. Interaction between cognate status and word frequency (Lg10WFc).
Note: $SUBTLEX_{US}$ word frequency is represented by centred log-transformed word frequency from the $SUBTLEX_{US}$ database (Lg10WFc).

Appendix E: Summary of multiple regression assumption tests conducted for the stepwise multiple regression analysis reported in section 7

The Kolmogorov-Smirnov test conducted on the residuals ($D(48) = .07$, $p = .2$) and the histogram of the standardised residuals indicated that the data contained approximately normally distributed errors. Tolerance statistics for all predictor variables were higher than .2, thus indicating that the assumption of no multicollinearity was met (Centred log of SUBTLEX-UK frequency, Tolerance = .51; Cognateness, Tolerance = 1; Centred log of SUBTLEX-UK frequency by Cognateness, Tolerance = .51). The loading of predictors on the smallest eigenvalue also indicated that the assumption of multicollinearity was met because most of the variance of only the cognateness variable was related to this eigenvalue (Centred log of SUBTLEX-UK frequency, variance proportion: .35; Cognateness, variance proportion: .54; Centred log of SUBTLEX-UK frequency by Cognateness, variance proportion: .34). The assumption of independent errors was also met (Durbin-Watson = 1.65). The scatterplot between the studentised residuals and the predicted standardised residuals indicates randomly scattered data points without any curvature or funnel shape; therefore, this scatterplot suggests that the assumptions of homoscedasticity and linearity were met. Finally, according to the power analysis reported in section 5, the optimal sample size for a regression analysis with six predictors, expected R^2 of .496, α level of .05 and 80% power is 21; therefore, 26 participants are more than sufficient for this regression analysis.

Table 1: Mean, median, minimum value, maximum value, range, standard deviation, skewness and kurtosis for predictor variables and the outcome variable

		Mean	Median	Min	Max	Range	SD	Skewness	Kurtosis
Cognates	SUBTLEX-UK English word form frequency per million words	16.78	9.55	0.95	77.28	76.34	20.18	1.9	3.26
	Log10(SUBTLEX-UK frequency)	0.93	0.98	-0.02	1.89	1.91	0.55	-0.01	-0.93
	Number of phonemes	8.81	8.5	3	13	10	2.53	-0.26	-0.16
	Number of letters	9.73	10	4	14	10	2.68	-0.12	-0.38
	Number of syllables	3.58	4	1	5	4	1.17	0.36	-0.74
	Item facility	0.95	0.94	0.82	1	0.18	0.05	-0.8	0.36
Non-cognates	SUBTLEX-UK English word form frequency per million words	17.56	9.39	1.04	80.47	79.44	20.95	1.85	3.03
	Log10(SUBTLEX-UK frequency)	0.95	0.97	0.02	1.91	1.89	0.55	-0.04	-0.98
	Number of phonemes	5.5	5.5	3	8	5	1.39	0.14	0.33
	Number of letters	7.23	7	4	11	7	1.86	-0.04	-0.57
	Number of syllables	2.04	2	1	4	3	0.72	0.64	1.15
	Item facility	0.89	0.93	0.65	1	0.35	0.11	-1.17	0.03
All words	SUBTLEX-UK English word form frequency per million words	17.17	9.55	0.95	80.47	79.52	20.37	1.89	2.72
	Log10(SUBTLEX-UK frequency)	0.94	0.98	-0.02	1.91	1.93	0.54	-0.03	-0.98
	Number of phonemes	7.15	7	3	13	10	2.62	0.52	-0.42
	Number of letters	8.48	8	4	14	10	2.61	0.33	-0.28
	Number of syllables	2.81	2.5	1	5	4	1.24	0.45	-0.83
	Item facility	0.92	0.94	0.65	1	0.35	0.09	-1.74	2.55

Table 2: Spearman's rho correlations among variables

		Log10(SUBTLEX-UK frequency)	Number of phonemes	Number of letters	Number of syllables
Cognates	Item facility	.42* (.03)	-.07 (.73)	-.26 (.2)	.01 (.95)
	Number of syllables	-.24 (.24)	.84* (<.001)	.78* (<.001)	
	Number of letters	-.3 (.14)	.94* (<.001)		
	Number of phonemes	-.3 (.14)			
Non-cognates	Item facility	.55* (.003)	-.07 (.74)	-.14 (.5)	-.03 (.88)
	Number of syllables	-.15 (.46)	.75* (<.001)	.75* (<.001)	
	Number of letters	-.26 (.2)	.78* (<.001)		
	Number of phonemes	-.27 (.19)			
All words	Item facility	.46* (.001)	.16 (.26)	-.01 (.92)	.2 (.17)
	Number of syllables	-.17 (.23)	.9* (<.001)	.84* (<.001)	
	Number of letters	-.25 (.08)	.9* (<.001)		
	Number of phonemes	-.21 (.146)			

Note. Two-tailed significance level is within parentheses. Significant correlations are starred.

Table 3: Best model of predictors of IF according to stepwise multiple regression analysis

	b	*SE b*	*B*	*t* value	*p*
Constant	.9	.01		68.04	<.001
Centred log10(SUBTLEX-UK frequency)	.12	.02	.77	4.97	<.001
Cognateness	.06	.02	.33	3.01	.004
Centred log10(SUBTLEX-UK frequency)*Cognateness	-.08	.03	-.38	-2.41	.02

Table 4: Importance of predictor variables according to squared semipartial correlations

Predictor	sr^2
Centred log10(SUBTLEX-UK frequency)	.30
Cognateness	.11
Centred log10(SUBTLEX-UK frequency) x cognateness	.07
Total	.48

Table 5: Best model of predictors of IF according to stepwise multiple regression analysis conducted on the data of participants who have staid less than three months in an English-speaking country

	b	*SE b*	*B*	*t* value	*p*
Constant	.88	.01		65.7	<.001
Centred log10(SUBTLEX-UK frequency)	.13	.02	.8	5.33	<.001
Cognateness	.07	.02	.4	3.69	.001
Centred log10(SUBTLEX-UK frequency)*Cognateness	-.09	.03	-.39	-2.6	.013

Table 6: Importance of predictor variables in the stepwise multiple regression analysis conducted on the data of participants who have staid less than three months in an English-speaking country according to squared semipartial correlations

Predictor	sr^2
Centred log10(SUBTLEX-UK frequency)	0.33
Cognateness	0.16
Centred log10(SUBTLEX-UK frequency) x cognateness	0.08
Total	0.57

Raquel Perez Urdaniz is an English as a Foreign Language teacher. She holds an MA from the University of Essex.

Sophia Skoufaki is a Lecturer at the University of Essex. She is an applied linguist mainly specialising in second language vocabulary learning and processing. Her other interests are reading comprehension and written discourse coherence.