Not just 'small potatoes': Knowledge of the idiomatic meanings of collocations

Marijana Macis and Norbert Schmitt
The University of Nottingham, UK

Abstract
This study investigated learner knowledge of the figurative meanings of 30 collocations that can be both literal and figurative. One hundred and seven Chilean Spanish-speaking university students of English were asked to complete a meaning-recall collocation test in which the target items were embedded in non-defining sentences. Results showed limited collocation knowledge, with a mean score of 33% correct. The study also examined the effects of frequency, semantic transparency, year at university, and everyday engagement with the second language (L2) outside the classroom on this collocation knowledge. Mixed-effects modelling indicated that there was no relationship between frequency and semantic transparency and the knowledge of the figurative meanings. However, a positive relationship was found between this knowledge and year at university, time spent in an English-speaking country, and time spent reading.

I Introduction

When is a collocation an idiom? When is an idiom a collocation? Take the example of *a piece of cake*, which appears both in collocation and idiom dictionaries. The reality is that these idiomatic[1] phrases exist in language, and what they are called depends largely on the research perspective of the namer. Scholars from the phraseological school (e.g. Howarth, 1996; Moon, 1998) will identify such 2-word combinations as figurative idioms, based on this school's emphasis on semantics and combinability criteria (e.g. *small potatoes* = 'something or someone insignificant'). Scholars from the frequency/ statistical school (e.g. Durrant, 2014; Sinclair, 1991) might extract the same

combinations from corpora, based on criteria of recurrence and statistical measures of co-occurrence, and call them collocations. However, the frequency school's highlighting of statistical metrics means that they have typically not considered semantics to any great extent (although, as exceptions to this, see Conrad & Biber, 2004; Hoey, 2005).

But while the question of terminology (figurative idiom or a collocation with a figurative meaning) might exercise academic minds, what is probably most important from a learner's perspective is that such phrases have an idiomatic meaning (i.e. cannot be understood from the combined meanings of the component words). Idiomatic phrases have been shown to be relatively difficult to master (e.g. Irujo, 1986), although much of this research has investigated longer, more prototypical idioms, e.g. *paper over the cracks*. There has been less research focusing on shorter idiomatic phrases, the kind that the frequency approach might throw up as collocations. In fact, we could find no research that extracted collocations based on frequency-based criteria, and then went on to investigate learners' knowledge of those that specifically carried figurative meanings. This study will explore second language (L2) learners' knowledge of idiomatic phrases of the type that the frequency approach would identify as collocations, and which factors facilitate the learning of such phrases.

II Background

1 The difficulty of lexical items with idiomatic meanings

There is considerable research literature on idiomatic language, and there seems to be a consensus that it is relatively challenging for second language learners; e.g. Celce-Murcia and Larsen-Freeman (1998, p. 39) refer to idioms as 'notoriously difficult'. For example, Irujo (1986) believes that idioms are problematic for the following reasons: they have non-literal meanings, but can also have literal counterparts that might confuse learners; idioms are often omitted from input to learners; most teaching materials do not include idioms; and even if idioms are known, learners often do not know when or how to use them appropriately. Likewise, phrasal verbs – which often have figurative meanings (*look up* = 'check in a reference source') – are often not particularly well-known (e.g. Dagut & Laufer, 1985). Given these difficulties, it is not surprising that learners often avoid using figurative formulaic sequences (e.g. Liao & Fukuya, 2004), even in informal spoken contexts where they might be more appropriate (Siyanova & Schmitt, 2007).

It is thought that native speakers generally learn and use the idiomatic meanings of idioms without much reference to their alternative literal meanings. After all, *to spill the beans* is much more often about spilling secrets than beans (Boers & Webb, 2015). Conversely, L2 learners are much more inclined to interpret idioms literally. For example, Martinez and Murphy (2011) demonstrate how learners often fail to recognize the idiomaticity of expressions and interpret them literally (e.g. *it's about time* interpreted as 'has a problem with time'). The difficulty of idioms applies to even relatively proficient learners, as international students at a British university often misunderstood the idioms used by lecturers there (Littlemore et al., 2011). This kind of misunderstanding can be especially difficult if the idioms have different underlying cultural basis, e.g. a *windbag* 'talks too much' in English would be *tong kosong* ('empty bowels') in Malay

(Charteris-Black, 2002). In sum, idiomatic language can be problematic for learners, from both productive and receptive standpoints.

2   Collocations with figurative meanings

Much of the idiom research to date has focused on longer idiomatic strings, e.g. *icing on the cake*. However, it has been noted that idiomatic language is not restricted to what might be considered 'prototypical' idioms, but can include phrases that are typically thought of as collocations (e.g. Boers & Webb, 2015). For example, Webb, Newton, and Chang (2013) define their target items as collocations, even though some of them can also be used figuratively (e.g. *pull strings*, *cut corners*). Similarly, Wolter and Gyllstad (2012) define the combination *bottom line* (= 'the important conclusion') as a collocation using the statistical approach, even though it has a figurative meaning. Furthermore, these idiomatic collocations might not be so rare. Macis and Schmitt (in press) analysed a small set of 54 collocations identified according to statistical criteria, and found that about 22% carried meanings that were figurative (usually in additional to a literal meaning; e.g. *top drawer* = 'something that is best of its class' and 'the uppermost drawer in a cabinet').

    The fact that collocations in general pose problems for L2 learners is well attested (e.g. Barfield, 2003; Wolter & Gyllstad, 2012). However, most studies into collocation knowledge have followed the statistical approach (for an overview of collocation studies, see Henriksen, 2013), which tends not to consider semantics as part of the selection process. As a result, the studies indicating learner problems with collocations have tended to use mostly collocations with literal meanings. There appears to be little research that focuses on learner knowledge of the figurative meanings of statistically-derived collocations, even though the idiom research reviewed above would suggest that such idiomatic meanings might be problematic.

3   Factors affecting lexical acquisition

Many factors have been shown to affect the acquisition of individual words, and it is likely that the same holds true for formulaic sequences, including idiomatic ones. For example, it has been suggested that L2 idioms that are congruent with first language (L1) idioms are easier to learn (Charteris-Black, 2002). Also, cross-cultural differences can have an effect. In Western societies, the heart is often associated with emotions (*to wear your heart on your sleeve*), while the mind is connected with reason (*to keep your head* in a frantic situation). But in Chinese, the mind is associated with *xin* ('heart'), so Chinese learners of English can find it difficult to understand English idioms based on the HEART = EMOTIONS metaphor (Hu & Fong, 2010). However, from the myriad factors than can affect language learning, it has been argued (from a usage-based perspective) that frequency and saliency have a particularly strong effect (e.g. Ellis, 2002), although it is not yet clear to what degree they also facilitate the learning of formulaic sequences.

a   The role of frequency. The proponents of usage-based models suggest that frequency plays the central role in language acquisition at all levels, from individual words to phrases (e.g. Ellis, 2002; Tomasello, 2003). Certainly frequency is one of the most robust

factors in vocabulary acquisition, and more frequent individual words are generally learned before less frequent ones (Ellis, 2002; Nation & Waring, 1997). There is some evidence that the same tendency also obtains with formulaic sequences. For example, Sonbul (2014) found that both native and non-native participants were sensitive to the frequency of adjective–noun collocations on an offline rating task, and that the sensitivity to frequency increased alongside the proficiency of the non-native speakers. Similarly, Durrant and Schmitt (2009) looked at the production by L2 learners of collocations of various frequencies extracted from the British National Corpus (BNC). They found that the learners used high frequency collocations extensively, but failed to use those that were less frequent but strongly associated (*MI* score > 3; Hunston, 2002).

But frequency might not explain everything. Reuterskiöld and Van Lancker Sidtis (2012) found relatively strong learning of idioms by young (9–14-year-old) L1 children from a single spoken exposure in a natural interactive context. Also, González Fernández and Schmitt (2015) found only a relatively weak link between frequency and productive collocation knowledge, and concluded that frequency cannot be used as the major predictor of collocation learning. This is closer to the results of Durrant's (2014) meta-analysis of nineteen collocation studies, which found that frequency correlated only moderately with collocation knowledge in the studies.

An important caveat is that most collocation studies to date have mainly used collocations with literal meanings (e.g. *strong man*, a 'free combination' in Howarth's terminology). So it is an open question to what degree the previous findings about frequency also apply to figurative collocations.

b   The role of semantic transparency. Some researchers claim that factors other than frequency may be more salient for L2 learners (Ellis, 2002; Wray, 2002). One of these might be semantic transparency. It has mostly been discussed in relation to idioms (e.g. Cowie 1981; Howarth 1996), but most formulaic sequences, including collocations, are opaque to some extent (Taylor, 2004). Figurative meanings have varying degrees of semantic transparency, and one might assume more transparent meanings are learned before less transparent meanings. If a learner knows the meaning of the two words making up a transparent (literal) collocation, then that collocation can be understood through decoding the constituents in their literal sense (*take the money*). The meaning of a semi-transparent collocation (*take a course*) is not decoded as easily as a literal counterpart, but is less salient than a non-transparent (figurative) collocation (*take sides*), which is very noticeable because it cannot be understood on the basis of its constituent parts. As a result, it has been argued that it is precisely the semi-transparent collocations that will cause problems for language learners (Nesselhauf, 2005). Indeed, Gyllstad and Wolter (2015) argue that the slower processing of collocations in their results was caused by the semi-transparent (figurative) nature of some of their collocations.

Moreover, some collocations can be 'deceptively transparent' (Boers et al., 2014) and this can cause deceptive comprehension, because L2 learners know the individual, usually very frequent general words, but are not familiar with these words in combination (Martinez & Murphy, 2011), especially when they carry a figurative meaning. Thus, overall, it would seem logical to assume that semantically less transparent collocations pose a greater challenge for L2 learners.

c   The role of engagement with the L2.   Language input is more than just about frequency. The quality of the input and interaction also matters. This suggests another factor in acquiring L2 collocations: learners' communicative engagement with their second language. Adolphs and Durrow (2004) found that there was a relationship between the quality of social integration and the amount of formulaic language produced in the speech of their two learners. Moreover, Schmitt and Redwood (2011) and González Fernández and Schmitt (2015) found a positive relationship between the amount of L2 engagement (e.g. reading, watching TV, and social networking in English) and knowledge of target phrasal verbs and the collocations respectively. These findings are consistent with Ellis's (2001) claim that frequent multi-word phrases, which fulfil a meaningful communicative function, will be more salient to learners, and therefore more likely to be learnt than those with less useful functions.

These positive findings about language engagement outside the classroom are interesting for at least two reasons. First, the merely-moderate relationship between frequency and collocation knowledge (Durrant, 2014; González Fernández & Schmitt, 2015) might be partially down to the fact that frequency counts are extracted from available corpora, and few, if any, of these are likely to be truly representative of the language to which any learner has been exposed (Durrant & Schmitt, 2010). It may be that the amount of language exposure of individual learners (e.g. the amount of language socialization, interaction, level of engagement) is a better indicator of the learning of collocations than corpus-based frequency.

A second interesting issue concerns second language acquisition (SLA) versus foreign language acquisition (FLA) contexts. Previously, SLA contexts (with high language availability) were considered much more advantageous because of the large amount of potential input outside the classroom. Also, frequency could be assumed to predict learning to a great degree, simply because learners could be expected to be exposed to language features in roughly the amounts which corpus frequency counts indicated. However, with the age of widespread electronic communication, FLA contexts may no longer be quite as impoverished as thought before. If learners are actively engaged outside the classroom with second language books, television, music, and social media, then this personal language engagement might be a better predictor of their L2 input (and thus acquisition) than corpus frequency in these FLA environments.

While there is some evidence that language use factors facilitate the knowledge of phrasal verbs (Schmitt & Redwood, 2011), and collocations with literal meanings (González Fernández & Schmitt, 2015), to our knowledge there is currently no research on how out-of-class language engagement affects learner knowledge of the figurative meanings of collocations that would be identified by the statistical approach.

This study will investigate learner knowledge of figurative collocations and what factors facilitate this knowledge, by asking the following questions:

1. How good is L2 learners' knowledge of the figurative meanings of collocations?
2. How do corpus frequency and semantic transparency relate to knowledge of the figurative meanings of collocations?
3. How strongly do education and language engagement factors relate to the knowledge of the figurative meanings of collocations?

III Methodology

1 Participants

The participants were 107 Chilean students of English from three Chilean universities. The age range was 18–36 year (mean = 21.80 years, SD = 3.01). There were 37 males and 70 females (34.6% and 65.4%, respectively). The majority had never visited an English-speaking country, and the length of stay for those that had ranged from 1 to 20 months. At the time of the test, they were all following partially English-medium undergraduate programs (English Language and Literature and English–Spanish Translation) in their respective universities. We recruited a relatively equal number of students from 1st through 4th years of study. Only Spanish speaking participants were selected for this study in order to control for the L1–L2 congruency effects. Unfortunately, we were unable to obtain proficiency measures for the participants, but all were successfully taking classes at university level in English.

2 Selection of target collocations

Our study focuses on the figurative meanings of collocations that have been identified through a statistical approach. There are many types of these collocations, but if we tried to systematically measure each one, the study would soon become unmanageable and difficult to interpret. We therefore limited our study to adjacent lexical collocations (or with only one intervening word, e.g. *break* a *leg*). We further restricted these to Verb+Noun and Adjective+Noun combinations, as these are the most researched types (Henriksen, 2013). Our statistically-based definition was that collocations needed to have a raw frequency in the COCA of *MI* ⩾ 3, following Hunston (2002).

   As there is no established collocations list, we extracted our target items from a range of different sources. First, we consulted different collocation dictionaries (e.g. the *Longman Collocations Dictionary and Thesaurus*, 2013; the *LTP Dictionary of Selected Collocations*, Hill & Lewis, 1998), searched for collocations on the internet, and looked into media like TV and radio. As Macis & Schmitt's (in press) results suggest that collocations with only figurative meanings are relatively rare, we focused our search on polysemous collocations with both literal and figurative interpretations ('duplex collocations' in Macis and Schmitt's terminology; 'figurative idioms' in Howarth's). Over 50 potential items were identified and included in a candidate pool. An additional native-speaking rater confirmed that the candidate items indeed had figurative meanings.

   Second, as our participants were Spanish speakers, we made sure that the target collocations were not directly translatable into Spanish. Two candidates were removed as a result of this check.

   Third, we needed collocations with a range of frequencies of figurative meaning (but with a minimum of 10 occurrences), because we wanted an indication of learners' knowledge of figurative meanings of collocations at different frequency levels. We therefore carried out a corpus analysis to determine the frequencies of both the figurative and the literal meanings for each potential target item. The analysis consisted of the first author reading 100 random concordance lines and tagging them for figurative or literal meaning, and then confirming the resulting figurative/literal ratios with a second random set

of 100 concordance lines. For the vast majority of items, the two ratios were very similar, and so the results from the two analyses were averaged. Unclear cases were referred to the second author, and sometimes to a third rater for resolution. In the vast majority of cases, the meanings were clear, and only a few items needed to be excluded because they were ambiguous.

In case of the majority of the target collocations (23/30), the figurative meaning was more frequent than the literal meaning. We felt this uneven distribution was acceptable, as collocations with dominant figurative meanings are more common than collocations with dominant literal meanings (Moon, 1998, p. 181). However, we also wanted to include collocations with dominant literal meanings, because in real life learners are exposed to all different types of collocations.

We chose the Corpus of Contemporary American English (COCA; Davies, 2008) as our reference corpus because it is large (>520 million words), is balanced across five different text types, and is regularly updated to include current changes in English. When calculating the frequency of collocations, we made the decision to search for each collocation as a word form, not lemma, as Sinclair (1991) argues for individual word form, because collocates are often different for the different word forms included in a lemma. An example of this is the idiom *take a hike*, which figuratively means 'go away'. Its lemmas *takes a hike*, *taking a hike*, *took a hike* and *taken a hike*, on the other hand, are typically used in their literal sense in the COCA. Finally, we selected the 30 collocations that produced the best range of figurative meaning frequencies (thereafter F frequency) and semantic transparency (see below) as our final target items; for the final list with both figurative and literal frequencies, see Appendix 1.

3   Test format

The research instrument consisted of two sections. In the first section, a meaning-recall test was designed to obtain a measure of the learners' knowledge of the figurative meanings of target collocations. The collocations were embedded in sentences, marked in bold and underlined. Below each sentence, there was space for the students to write the definition.

Because of her personality, she is known as the **queen bee**.

_____

The participants were free to answer in either English or Spanish, to ensure that any students with lower proficiency levels in English could still demonstrate his or her knowledge. The consent form and the instructions were written in both English and Spanish, while the test presenting the target was written in English only. Meaning recall tests measure learners' ability to recognize and understand an L2 form when it is presented. In this case, the form was written, so the level of knowledge would be analogous to that necessary to understand the collocation in a written text, i.e. receptive knowledge.

The second section was developed to collect information about the degree of engagement with the L2. It consisted of a questionnaire with (1) biodata questions concerning

gender, age, year at university, and whether they had travelled abroad, and (2) questions related to the participants' English language use outside the classroom (reading, watching TV, listening to music, and social media). The complete instrument is available as Supplementary Material online.

4   Test piloting

Two pilotings were conducted to check the validity of the test for the purposes of the research. First, we wanted to ensure that the figurative meanings of the collocations could not be guessed solely on the basis of the context sentences. Ten native speakers of English (9 females and 1 male) were given the context sentences with the collocations deleted to see if they could guess the missing phrases. The results showed that the sentence contexts were essentially not guessable by the native respondents, and so were highly unlikely to be guessable by our nonnative main study participants. We can thus be confident that any correct answers were not the result of contextual guessing.

The second piloting was carried out to determine whether the test was suitable for non-native speakers. The intact test (30 sentences with the target items inserted) was administered to 10 non-native speakers (6 males and 4 females, all of them postgraduate students at a British university). The results indicated weak collocation knowledge in general. Follow-up interviews showed that the non-natives did not find anything misleading on the test, and that they tried to guess the meaning of the target collocations whenever they were not sure about it. If they could not guess, they would leave the space blank. These results indicated that the test was ready to be used in the main study.

5   Transparency task

Semantic transparency was one of the factors we wished to investigate, and so we needed to determine the transparency of the target collocations. Transparency is not a notion that can be quantified in absolute terms. Rather, it resides on a continuum and so must be quantified by subjective ratings. Thus, a transparency task in our study was conducted to see how easy or difficult it was to guess the figurative meaning of the target collocations based on their literal meanings. As the 'reliability [of transparency ratings] is enhanced when the estimates of several raters are pooled' (Boers & Webb, 2015, p. 383), 18 native speakers of Spanish (13 Chilean, 3 Mexican and 2 Spanish participants) took part in the task, all with university degrees. They first received a detailed explanation of collocations and transparency. They then were provided with both the literal and figurative meaning of each collocation and asked a) to state whether they knew the figurative meaning and b) to rate each figurative meaning on a scale from 1 to 4: 1 being very difficult to guess (very opaque) and 4 being very easy to guess (very transparent). As desired, the collocations showed good diversity in their transparency (min = 1.44, max = 3.56, SD = .650).

6   Administration and data analysis

The test was administered in Chile. No time limits were set, but most students finished in about 30–45 minutes. The maximum score for each test was 30, based on one point per

correct figurative meaning. Accurate spelling and correct inflections were not required for a collocation to be marked as correct, as long as the meaning definitions were comprehensible and clear. The first author scored the test and 30% of the sample was marked by a second rater. There was a 94.5% agreement between the raters.

To address the second and third research questions, mixed-effects models were chosen because they allow for the inclusion of both subject and item as random effects. This allows the researcher to account for individual differences in subjects (e.g. first-year versus second-year students) as well as in items. It also eliminates the need for separate analyses with participants as a random variable and items as a random variable (F1 and F2 analyses). The data was thus submitted to a mixed effects modelling analysis (R package lmerTest, Version 2.0–11; Kuznetsova et al., 2014) to determine how the various factors related to the knowledge of figurative meanings of collocations.

IV   Results

1   How good is L2 learners' knowledge of the figurative meanings of collocations?

To answer the first research question, we calculated the mean of correct answers. Table 1 shows the descriptive statistics for our Spanish-speaking university students of English as a foreign language (EFL). As we can see, the mean of correct answers was 33.02% (9.91/30). The standard deviation (4.49) and the range of correct answers (0–21) indicate that there was a relatively small amount of variation across the sample. This is illustrated in Figure 1, which shows the distribution of the test scores.

Figure 1 shows that the vast majority (81 participants, 75.7%) scored between 5 and 14. There were no scores beyond 21, so it seems that the test was equally hard for everybody. Also, there were two students who scored 0, which is another indication of how difficult it may be to learn the figurative meanings of collocations.

2   How do corpus frequency and semantic transparency relate to knowledge of the figurative meanings of collocations?

Results were analysed using an omnibus linear mixed effects model using the lme4 package (Version 1.1-10, Bates, et al., 2014) in R (Version 3.2.2, R Core Team, 2015). The model development procedure was conducted in the following way. First, we extrapolated the F frequencies in order for them to be representative of the total number of occurrences of the figurative meanings in the COCA. Second, we log transformed the total F frequencies to reduce skewing as the data had a wide range (from 5 to 797). Then, we centred all continuous variables, e.g. transparency. Because our independent variable (knowledge of figurative meanings) was binary, we used a generalized linear model with binomial regression.

A model was created that included two variables that have been shown to be important in figurative meaning: F frequency and transparency. The first model included fixed effects of F frequency and transparency, along with random effects of subject and item.

Table 1. Descriptive statistics of the participants' test scores (maximum = 30).

	n	Minimum	Percentage	Maximum	Percentage	Mean	SD	Percentage
Participants' scores	107	0	0.0	21	70.0	9.91	4.49	33.02
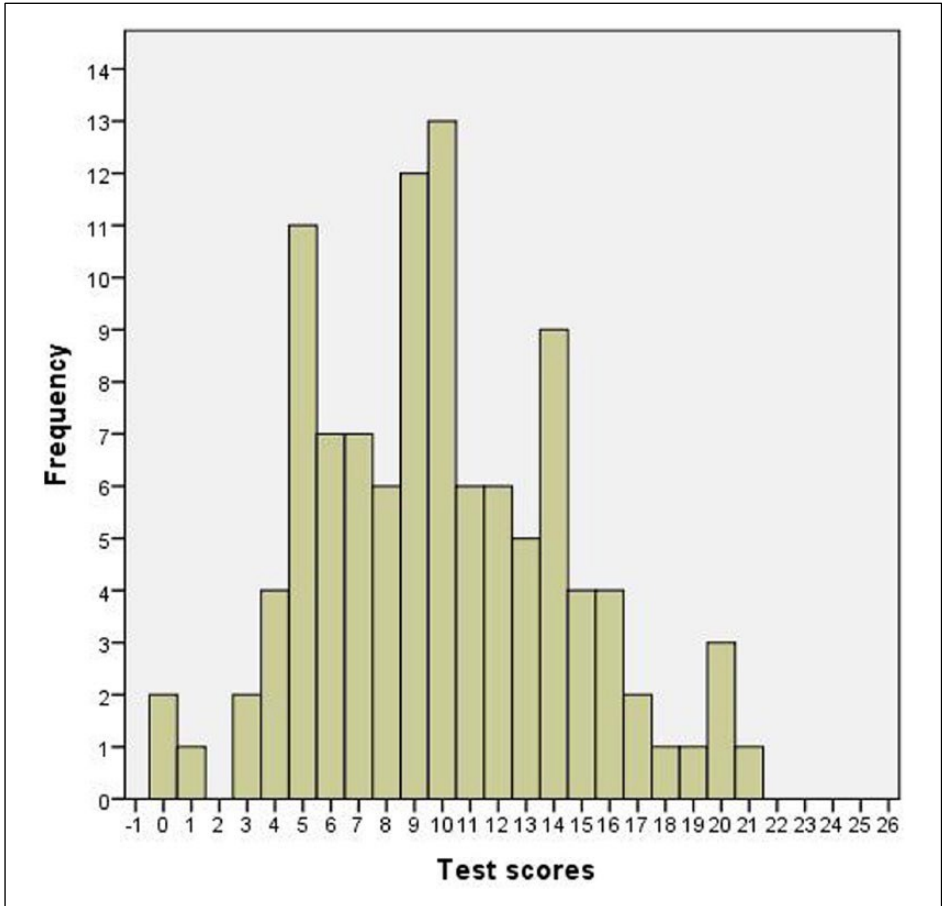
Figure 1. Distribution of test scores.

The results of the model (see Supplementary Material: Model 1) showed that neither of the variables was significant. We also constructed a model looking at the interaction of these two fixed effects. Since neither model showed significant effects or a significant interaction, these variables were discounted from future models. Thus neither F frequency nor semantic transparency were related to knowledge of the figurative meanings of the collocations in our study.

3 How strongly do education and language engagement factors relate to the knowledge of the figurative meanings of collocations?

In order to answer the third research question, we built a comprehensive model that included the other variables, i.e. year at university, time spent in an English-speaking country, and time spent reading, watching TV, listening to music, and social networking in English (see Supplementary Material: Model 2). We used a backwards stepwise procedure to eliminate variables that did not significantly improve the overall fit of the model. The process involved eliminating the variable with the lowest z-score and then refitting the model. This procedure continued until all insignificant variables were removed. We first eliminated TV, music, and social networking. We used explicit model comparisons to confirm that removal of these variables did not significantly reduce the fit of the model (i.e. the more complex model was not a significant improvement). As a result, the final model[2] (Table 2) included knowledge of figurative collocations as the independent variable, with year at university, time spent in an English-speaking country, and time spent reading in English as significant covariates.

We then compared the initial comprehensive model (Model 2 with all variables) and the final model, and there was no statistically significant difference between the two. Finally, we tried adding F frequency and transparency back into the final model to check whether they were significant when the three other significant factors were included. No significant effects were seen for either variable, either in random intercept only models, or in models with by item random slopes for the effects of F frequency and transparency.

To sum up, the omnibus analysis shows clear effects of year at university, time spent abroad, and time spent reading in English. However, there was no effect for other factors, including F frequency and transparency.

V Discussion

Although we were not able to obtain a proficiency measure for our participants, they were doing university degrees in partially English-medium courses at respected universities, which would imply a relatively high level of English. Nevertheless, our study found that they knew only about 33% of the target figurative collocation meanings. But should we interpret this as a relatively good or weak performance? In this case, terminology matters. If one wishes to view the target phrases as 'figurative idioms', then they might compare these results to other results from idiom studies. These usually show quite low levels of mastery – e.g. virtually nothing at lower proficiencies up to around 22% at Certificate of Proficiency (CPE) level (McGavigan, 2009, reported in Milton, 2009, pp. 151–159) – and so our results might look quite encouraging in comparison. Alternatively, results of 'collocation' studies (usually statistically-derived) often show better knowledge (e.g. 78%–82%; Gyllstad, 2009). Compared to these figures, the results are not so impressive. Perhaps the most straightforward way of viewing the results is that our relatively advanced students still struggled with the figurative meanings of collocations, suggesting these meanings are a problematic feature for learners. In addition, it is also important to note that our participants were tested at a receptive meaning recall level of mastery, and use of a productive test (form recall) would probably have yielded even lower scores.

Table 2. Mixed effects modelling of factors affecting knowledge of figurative collocations.

Fixed effect	Estimate	Standard error	z-value	p-value
(Intercept)	−2.041	0.311	−6.568	5.09e-11***
Year at university	0.297	0.069	4.319	1.57e-05***
Visit Abroad	0.058	0.027	2.152	0.031*
Reading	0.020	0.007	2.881	0.004**
Random effects	Variance			
Subject	0.4506			
Item	1.7118			

Notes. Significance values are estimated by the R package lmerTest (version 2.0–11; Kuznetsova et al., 2014); *** $p < .001$, ** $p < .01$, * $p < .05$.

These results are congruent with previous findings showing the lack of collocational knowledge in general (e.g. Barfield 2003; Wolter & Gyllstad, 2012). Those studies usually included only or mostly literal collocations, and we found that collocations with figurative meanings were also difficult for our learners. This was not unexpected, given the problems learners have in general with idiomatic formulaic sequences (i.e. idioms, phrasal verbs) (e.g. Irujo, 1986), and our study gives initial evidence that this difficulty also extends to the figurative meanings of collocations. Collocations are important for language use, but seem to be problematic for learners, regardless of whether they carry literal or figurative meanings.

Nevertheless, our participants did know some of the figurative meanings, and a relatively few (11/107: 10%) knew one-half or more. So what factors related to this knowledge? Frequency is usually the leading candidate, but in this case it showed no statistically reliable effect. Furthermore, it did not seem to make much difference which meaning sense (figurative or literal) was more frequent for the collocations. In fact, if anything, there was a slight advantage in knowledge of figurative meanings for collocations in which the literal meaning was more frequent, e.g. *brick wall* was the 11th best known collocation, even though its literal meaning was more frequent. These results are in contrast to the many studies that show a robust effect of frequency on the acquisition of individual vocabulary items (e.g. Ellis, 2002; Tomasello, 2003). We feel three explanations are possible. First, it may be that formulaic language does not follow frequency nearly as strongly as individual words. Schmitt and Redwood (2011) found relatively weak relationships between frequency and knowledge for phrasal verbs (receptive: $r^2 = 9$–13%; productive: $r^2 = 18$–20%) and González Fernández and Schmitt (2015) found similar low results for collocations (productive: $r^2 = 20\%$). Our receptive meaning recall test scores did not correlate to figurative frequency at all (Spearman, p = .469, *ns*). It is unclear why frequency should relate more weakly to formulaic sequences than individual words, but one might speculate that formulaic sequences are less noticeable in language because learners tend to process them word by word (Wray, 2002). This might negate the effect of sequence frequency somewhat.

Another explanation revolves around Durrant and Schmitt's (2010) idea that corpus frequency counts may not correspond very well to the amount of actual input learners receive in their own environments. While these counts would hopefully be fairly

representative in environments of English as a second language (ESL) where the L2 is freely and widely available, they may not hold in EFL environments where the main input is driven by materials writers and teachers. The input from instructed EFL environments may not relate strongly to corpus frequency simply because vocabulary selection tends to be relatively unprincipled and opportunistic in both written materials (Schmitt & Schmitt, 2014) and teacher talk (Horst, 2010). In EFL situations, learners get less exposure overall, and it can be quite different from the frequency profiles in more widespread language as represented by corpus frequency counts.

A third explanation is the degree of frequency. On average, individual words are more frequent than formulaic sequences, including collocations. It might be that the relatively higher frequency of individual words (compared to the relative frequency of most collocations) is sufficient to drive acquisition in a usage-based (Ellis, 2002) manner. Conversely, the relatively lower frequency of most collocations may not be enough to push acquisition in a systematic manner that shows up in statistics.

Similarly, the relative semantic transparency of the figurative meanings did not seem to matter either. Some very transparent collocations, like *queen bee* were known by almost 75% of the participants. Conversely, other collocations of similar transparency were answered correctly by a much lower percentage of the participants (e.g. *rainy day*: 7.48%). This may seem somewhat surprising, as degree of semantic transparency is often seen as factor affecting difficulty (e.g. Martinez & Murphy, 2011). Our non-significant result might be explained by the fact that people's intuitions about transparency are based on the degree to which they can see the relationship between individual components and their figurative referents. This is personal and subjective, and can vary a great deal from person to person. For example, in case of collocations such *hit the road* and *small potatoes*, there was a range of answers on the transparency task. A much larger norming sample might even out this variability to some extent (e.g. 100s or 1,000s of respondents as in word association norms (e.g. Nelson et al., 2015), but perhaps semantic transparency is simply too subjective to relate reliably to collocation knowledge when exposed to the precision of statistical analyses. Along these lines, Tabossi et al. (2008) concluded that people do not have clear and systematic intuitions about semantic compositionality of idioms. In future research, it might be interesting to explore this question in a more personal manner by using interviews. They could explore what kind of mental associations people make regarding the relationship between individual components of a collocation and its figurative meaning, and how this affects resulting knowledge.

So F frequency and semantic transparency did not relate to figurative collocation knowledge, which leads us to amount of education. 'Year at university' related to knowledge and we could speculate that the more proficient learners become (assuming that more years at a (partially) English-medium university equates to higher proficiency), the more collocation knowledge they will have. Indeed, proficiency generally has been shown to have a facilitative effect (e.g. Gyllstad, 2007; Eyckmans, 2009). But even if we do not allow the assumption about higher proficiency, our results demonstrate that more years of university education leads to better knowledge of figurative collocations.

Finally, language use factors showed mixed results, with only time spent in an English-speaking country and the hours of L2 reading per week having an effect. The 'time spent abroad' result is consistent with the existing research on facilitative nature of

the L2 environment (e.g. González Fernández & Schmitt, 2015; Milton, 2009). It is generally known that extensive exposure to the L2 environment through social and cultural adaptation and an on-going contact with local native speakers can lead to students' increase of the use of formulaic sequences in their L2 production and their overall language proficiency (e.g. Adolphs & Durrow, 2004). This is in line with Wray's (2002) view that naturalistic settings differ from the language in a classroom, since the former contains more formulaic sequences that fulfil social and communicative functions. It might explain why the participants who spent some time in an L2 country obtained better scores than those who never had such an opportunity.

Research has consistently shown that reading facilitates both vocabulary knowledge (e.g. Horst et al., 1998) and overall language proficiency (Renandya et al., 1999). Our results show that reading also facilitates the acquisition of figurative meanings of collocations. Furthermore, this is congruent with findings from studies focusing on other types of formulaic sequence. Schmitt and Redwood (2011) and González Fernández and Schmitt (2015) found that the amount of reading had a positive relationship with knowledge of phrasal verbs and collocations, respectively. Thus, our results suggest not only that reading outside the classroom facilitates the acquisition of the figurative meanings of collocations, but also adds to the converging evidence that it benefits the learning of formulaic language in general.

In the literature review, we speculated whether the ever-increasing electronic access to an L2 might make EFL contexts more like ESL ones. This seems to be the case with reading, but none of the other possible language inputs reliably related to collocation knowledge (watching TV, listening to music and social networking). In case of the social media, the result is somewhat surprising, as this is the type of language use which should make language interesting and meaning-based. González Fernández and Schmitt (2015) found a significant relationship between social networking and collocation knowledge, but perhaps figurative meanings are less salient and go unnoticed.

Time spent watching TV did not have any effect either. Even though learners can be exposed to a wide range of vocabulary through films and TV (Rodgers & Webb, 2011), including the figurative meanings of collocations, perhaps they simply do not notice them as long as they can understand the general message. This speculation is supported by research that shows the effectiveness of subtitles in facilitating learning (e.g. Neuman & Koskinen, 1992), presumably by promoting noticing of the target items and showing their written forms. The same explanation may apply to listening to music: it does not require as much attention and/or concentration as reading, unless it is accompanied by other more conscious tasks (Beasley et al., 2008). Learners might also listen more to the tune, than the words/phrases.

All of the findings must be interpreted in light of the inevitable limitations of our study. The amount of L2 engagement of the participants was assessed via self-report questionnaires, and thus could be prone to slight underestimation or overestimation. The participants were a fairly homogeneous group, all university students with only one L1: Spanish, and therefore further research involving more diverse samples of L2 populations will be needed in order to make more robust generalizations. We were unfortunately not able to obtain a measure of the participants' language proficiency. We also only studied Verb+Noun and Adjective+Noun lexical collocations, and other types may

be known to greater or lesser degrees. Our results pertain to receptive meaning recall knowledge, and presumably productive form recall knowledge would be weaker, although this is a matter for further research.

VI   Implications

Whatever approach one uses to define collocations, some of the combinations identified will have figurative meanings. But the idea that collocations can be polysemous and that some of those meanings can be figurative still seems to be very novel when it comes to pedagogy. We suspect most teachers have never thought about these possibilities, but if Macis and Schmitt's (in press) preliminary calculations hold up, then figurative meanings of collocations are not just a peripheral phenomenon that can be ignored. Collocations in general are a widespread and important feature for language, and figurative meanings may well make up a substantial percentage of the total occurrences (i.e. perhaps one-fifth to one-quarter).

Our results suggest that figurative meanings of collocations will not be learned very comprehensively if left to themselves. It is a question for future research whether the more input-rich ESL environments will suffice for adequate incidental learning to occur without instruction. But at least in EFL environments like we studied, there does not seem to be enough input to ensure that a high percentage of these meanings will be learned. So what can be done? Spending time in an L2 country is one solution that seems to have an effect (as it probably would with most language features), but is obviously impractical in most situations. Encouraging students to engage with the L2 outside the classroom does not always work either, as watching TV, listening to music, and using social networking all proved ineffective in this case.

It seems that figurative meanings might not be learned very efficiently unless some attention is given to them. This seems to suggest good old-fashioned educational values: going to school and reading a lot. There is no reason to think that our target collocations were explicitly taught to our participants (at least not in any systematic way), but there seems to be something in the instructed environment which is facilitative, as students in higher university years knew more collocations. This suggests that figurative meanings of collocations would benefit from explicit attention in the classroom. As a start, teachers and learners need to be aware that not all collocations have literal meanings, and that some will carry figurative meanings.

Reading seems to be a powerful facilitator for language learning, with extensive reading being tied to improved language learning overall (e.g. Day & Bamford, 1998), and this beneficial effect seems to work with the figurative meanings of collocations to some extent as well. In fact, there is converging evidence that reading is a consistent moderate predictor of the acquisition of several categories of formulaic language (phrasal verbs: Schmitt & Redwood, 2011; collocations in general: González Fernández & Schmitt, 2015; and the figurative meanings of collocations as indicated in this study).

This incidental learning from reading might be even stronger if it were not for the relative infrequency of idiomatic phrases in the texts/listening. Boers and Lindstromberg (2009) found that, although idioms were common as a class, individual idioms are unlikely to appear frequently enough to facilitate their acquisition. They identified 42 occurrences of verb–noun collocations (e.g. *commit suicide*) in

120 pages of the novel *Beneath the bleeding*, but 35 (83%) occurred only once, and the best case was three repetitions (for only three collocations). They looked at longer idioms, but it may well be that the shorter figurative collocations are no more frequent, which would limit their availability for incidental learning. One solution to this problem be might be modified materials where idiomatic language is 'seeded' into the text at higher rates of recurrence.

There is also growing evidence from the area of cognitive linguistics (CL) that explicit instruction using underlying metaphors (e.g. ANGER IS HEAT) can help learners understand idiomatic language, as has been shown by series of studies by Boers and colleagues (e.g. Boers, Eyckmans, & Stengers, 2007). In an overview, Boers (2013) reviewed over 20 published studies, and found that CL-informed approaches were usually more effective than comparison treatments. However, it might be that this approach works better for longer idiomatic phrases (*being hot under the collar*, *adding fuel to the fire*) than for the shorter two-word combinations studied here, and this is an issue for future research.

CL-informed instruction might be useful, but the problem remains in determining which collocations to teach. Frequency is the normal way of grading vocabulary for difficulty/importance for teaching, and this works reasonably well for high frequency individual words. But formulaic language does not seem to follow a frequency profile to nearly the same extent, and so frequency may not be such a useful guideline for selecting which collocation items to focus upon. Without an obvious way to select collocations for explicit instruction, perhaps the way forward is to focus the explicit attention on strategy instruction and how to use dictionaries well to look up collocation meaning, and how to inference from context more effectively.

Notes

1. In this article we use the terms 'idiomatic' and 'figurative' interchangeably to refer to the non-compositionality of phrases.
2. We also created a model (Model 3) with all the variables included, but the outcome was the same, i.e. frequency and transparency were not significant, and the only significant covariates that came out of the model selection process were the ones listed in Table 2.

References

Adolphs, S., & Durrow, V. (2004). Social-cultural integration and the development of formulaic sequences. In: N. Schmitt (Ed.), *Formulaic sequences* (pp. 107–126). Amsterdam/ Philadelphia, PA: John Benjamins.

Barfield, A. (2003). *Collocation recognition and production: Research insights*. Tokyo: Chuo University.

Bates, D., Maechler, M., Bolker, B., et al. (2014). lme4: Linear mixed-effects model using Eigen and S4, R package version 1.1–10. Available at: http://cran.r-project.org/web/packages/lme4/ (accessed April 2016).

Beasley, R., Chuang, Y., & Liao, C.C. (2008). American music immersion: Influencing factors and its impact on Taiwanese EFL learners engaged in web-based, multimedia music study. *Journal of Educational Multimedia and Hypermedia*, *17*, 27–42.

Boers, F. (2013) Cognitive linguistic approaches to second language vocabulary: Assessment and integration. *Language Teaching*, *46*, 208–224.

Boers, F., & Lindstromberg, S. (2009). *Optimizing a lexical approach to instructed second language acquisition*. Basingstoke, UK: Palgrave Macmillan.

Boers, F., & Webb, S. (2015). Gauging the semantic transparency of idioms: Do natives and learners see eye to eye? In: R. Heredia, & A. Cieslicka (Eds.), *Bilingual figurative language processing* (pp. 368–92). Cambridge: Cambridge University Press.

Boers, F., Eyckmans, J., & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: More than mere mnemonics? *Language Teaching Research*, *11*, 43–62.

Boers, F., Lindstromberg, S., & Eyckmans, J. (2014). Some explanations for the slow acquisition of L2 collocations. *Vigo International Journal of Applied Linguistics*, *11*, 41–61.

Celce-Murcia, M., & Larsen-Freeman, D. (1998). *The grammar book: An ESL/EFL teacher's course*. 2nd edition. Boston, MA: Heinle ELT.

Charteris-Black, J. (2002). Second language figurative proficiency: A comparative study of Malay and English. *Applied Linguistics*, *23*, 104–133.

Conrad, S., & Biber, D. (2004). The frequency and use of lexical bundles in conversations and academic prose. In: W. Teubert, & M. Mahlberg (Editors of the thematic part), The corpus approach to lexicography. *Lexicographica: International Annual for Lexicography*, 20, 56–71.

Cowie, A.P. (1981). The treatment of collocations and idioms in learners' dictionaries. *Applied Linguistics*, *2*, 223–235.

Dagut, M.B., & Laufer, B. (1985). Avoidance of phrasal verbs by English learners, speakers of Hebrew – a case for contrastive analysis. *Studies in Second Language Acquisition*, *7*, 73–79.

Davies, M. (2008). The Corpus of Contemporary American English: 450 million words, 1990–present. Available at: http://corpus.byu.edu/coca/ (accessed April 2016).

Day, R.R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: Cambridge University Press.

Durrant, P. (2014). Corpus frequency and second language learners' knowledge of collocations: A meta-analysis. *International Journal of Corpus Linguistics*, *19*, 443–477.

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics*, *47*, 157–177.

Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research*, *26*, 163–188.

Ellis, N.C. (2001). Memory for language. In: P. Robinson (Ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press.

Ellis, N.C. (2002). Frequency effects in language processing and acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, *24*, 143–188.

Eyckmans, J. (2009). Toward an assessment of learners' receptive and productive syntagmatic knowledge. In: A. Barfield, & H. Gyllstad (Eds.), *Researching collocations in another language* (pp. 139–152). New York: Palgrave Macmillan.

González Fernández, B., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have? The effects of frequency and amount of exposure. *ITL International Journal of Applied Linguistics*, *166*, 94–126.

Gyllstad, H. (2007). Testing English collocations: Developing receptive tests for use with advanced Swedish learners. Unpublished doctoral dissertation, Centre for Languages and Literature, Lund University, Sweden. Available at: http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=599011&fileOId=2172422 (accessed April 2016).

Gyllstad, H. (2009). Designing and evaluating tests of receptive collocation knowledge: COLLEX and COLLMATCH. In: A. Barfield, & H. Gyllstad (Eds.) *Researching collocations in another language* (pp. 153–170). Basingstoke: Palgrave Macmillan.

Gyllstad, H., & Wolter, B. (2015). Collocational processing in the light of a phraseological continuum model: Does semantic transparency matter? *Language Learning*. Epub ahead of print 6 October 2015. DOI: 10.1111/lang.12143.

Henriksen, B. (2013). Research on L2 learners' collocational competence and development: A progress report. In: Bardel, C., Lindqvist, C., & Laufer, B. (Eds.), *L2 vocabulary acquisition, knowledge, and use* (pp. 29–56). EuroSLA Monograph Series 2. EuroSLA.

Hill, J., & Lewis, M. (1998). *LTP dictionary of selected collocations*. Hove: Language Teaching Publications (LTP).

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London: Routledge.

Horst, M. (2010). How well does teacher talk support incidental vocabulary acquisition? *Reading in a Foreign Language*, *22*, 161–180.

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, *11*, 207–223.

Howarth, P. (1996). *Phraseology in English academic writing: Some implications for language learning and dictionary making*. Lexicographica Series Maior, volume 75. Tübingen: Niemeyer.

Hu, Y.-H., & Fong, Y.-Y. (2010). Obstacles in CM-guided L2 idiom interpretation. In: S. De Knop, F. Boers, & T. De Rycker (Eds.), *Fostering language teaching efficiency through cognitive linguistics* (pp. 293–316). Berlin: Mouton de Gruyter.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Irujo, S. (1986). A piece of cake: Learning and teaching idioms. *ELT Journal*, *40*, 236–242.

Kuznetsova, A., Brockhoff, P.B., & Christensen, R.H.B. (2014). The R package lmerTest (version 2.0–11). Retrieved April 2016 from https://cran.r-project.org/src/contrib/Archive/lmerTest/.

Liao, Y., & Fukuya, Y.J. (2004). Avoidance of phrasal verbs: The case of Chinese learners of English. *Language Learning*, *54*, 193–226.

Littlemore, J., Trautman Chen, P., Koester, A., & Barnden, J. (2011). Difficulties in metaphor comprehension faced by international students whose first language is not English. *Applied Linguistics*, *32*, 408–429.

Longman Collocations Dictionary and Thesaurus (2013). Harlow, UK: Pearson Education Limited.

Macis, M., & Schmitt, N. (in press). The figurative and polysemous nature of collocations and their place in ELT. *ELTJ*.

Martinez, R., & Murphy, V. (2011). Effect of frequency and idiomaticity in second language reading comprehension. *TESOL Quarterly*, *45*, 267–290.

Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.

Moon, R. (1998). *Fixed expressions and idioms in English: A corpus-based approach*. Oxford: Oxford University Press.

Nation, I.S.P., & Waring, R. (1997). Vocabulary size, text coverage, and word lists. In: N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge: Cambridge University Press.

Nelson, D.L., McEvoy, C.L., & Schreiber, T.A. (2015) The University of South Florida Free Association Norms. Available at: http://w3.usf.edu/FreeAssociation (accessed April 2016).

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: John Benjamins.

Neuman, S.B., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, *27*, 95–106.

Renandya, W.A., Rajan, B.R.S., & Jacobs, G.M. (1999). Extensive reading with adult learners of English as a second language. *RELC Journal*, *30*, 39–60.

Reuterskiöld, C., & Van Lancker Sidtis, D. (2012). Retention of idioms following one-time exposure. *Child Language Teaching and Therapy*, *29*, 219–231.

Rodgers, M.P.H., & Webb, S. (2011). Narrow viewing: The vocabulary in related television programs. *TESOL Quarterly*, *45*, 689–717.

Schmitt, N., & Redwood, S. (2011). Learner knowledge of phrasal verbs: a corpus-informed study. In: F. Meunier, S. De Cock, G. Gilquin, & M. Paquot (Eds.), *A taste for corpora: In honour of Sylviane Granger* (pp. 173–209). Amsterdam: John Benjamins.

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

Sinclair, J. (1991). *Corpus, concordance, collocation: Describing English language*. Oxford: Oxford University Press.

Siyanova, A., & Schmitt, N. (2007). Native and nonnative use of multi-word vs. one-word verbs. *International Review of Applied Linguistics*, *45*, 119–139.

Sonbul, S. (2014). Fatal mistake, awful mistake, or extreme mistake? Frequency effects on offline/on-line collocational processing. *Bilingualism: Language and Cognition*, *18*, 419–437.

Tabossi, P., Fanari, R., & Wolf, K. (2008). Processing idiomatic expressions: Effects of semantic compositionality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 313–327.

Taylor, J.R. (2004). Why construction grammar is radical. *Annual Review of Cognitive Linguistics*, *2*, 321–348.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, *63*, 91–120.

Wolter, B., & Gyllstad, H. (2012). Frequency of input and L2 collocational processing. *Studies in Second Language Acquisition*, *35*, 451–482.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Appendix 1. List of target items with frequencies of figurative and literal meanings.

	Duplex collocation	Raw frequency in the COCA*	Extrapolated frequency of figurative meanings	Extrapolated frequency of literal meanings
1	Red tape	801	797	4
2	Red flag	691	605	86
3	Hit the road	550	536	14
4	Free ride	412	342	70
5	Hold one's breath	620	316	304
6	White collar	364	278	86
7	Blue ribbon	391	270	121
8	Dark horse	255	251	4
9	Brick wall	674	243	431
10	Rainy day	461	177	284
11	Old hat	178	153	25
12	Hold the line	180	152	28
13	Cold feet	219	151	68
14	Thick skin	152	126	26
15	Acid test	121	112	9
16	Small potatoes	124	107	17
17	Old hand	121	103	18
18	Fat cat	110	91	19
19	Take a hike	149	87	62
20	Sore spot	117	84	33
21	Hold water	150	72	78
22	Queen bee	99	72	27
23	Break a leg	114	49	65
24	Drop the ball	63	47	16
25	Big wheel	93	41	52
36	Top drawer	214	28	186
27	Dead duck	33	27	6
28	Hit the roof	41	22	19
29	Fancy pants	18	14	4
30	Bend (one's) knee	24	5	19

Note. COCA = Corpus of Contemporary American English (Davies, 2008).