

## Minding the Gap in Vocabulary Knowledge: Incidental Focus on Collocation through Reading

**Thamer Alharthi**

Department of European Languages and Literature  
King Abdulaziz University  
Jeddah, Saudi Arabia

### Abstract

The present study seeks to contribute to our knowledge of the effectiveness of reading in the incidental learning of collocations. The study also addresses the question whether out-of-class exposure such as watching TV, listening to radio or music, reading English books and using social media plays a significant role in the learning of collocations. The research participants were 46 Arabic-speaking young adult EFL learners. They were asked to read a modified text containing 10 pseudo-word collocations and to verify that all were unfamiliar to them. One week later, they read a text containing the actual 21 target collocations, which had been selected on the basis of appearing in instructional materials and a reference corpus, as well as of frequency. Participants were then asked to complete a meaning-recall cloze test in the form of a gap-filling task in which the 21 target collocations were embedded in sentences. Subsequently they were administered a self-report survey about any incidental exposure to English. The analysis of the quantitative results revealed that the target collocations can be learned incidentally through reading although the level of mastery was limited. The survey data showed a positive correlation between the learners' knowledge of collocations and activities such as watching TV, listening to radio and reading English books. The study also confirmed that collocations are particularly difficult for adult EFL learners and is thus an aspect of vocabulary knowledge in need of further empirical investigations.

*Keywords:* collocation, EFL, incidental learning, pseudo-words, reading

**Cite as:** Alharthi, T. (2018). Minding the Gap in Vocabulary Knowledge: Incidental Focus on Collocation through Reading. *Arab World English Journal*, 9 (2).

DOI: <https://dx.doi.org/10.24093/awej/vol9no2.1>

## Introduction

It is well established that vocabulary knowledge and its learning play a vital role in second language (L2) development. There is broad consensus that there are many aspects of a word to know, as Nation (1990, 2001) points out, who to date has provided the best specification of such word knowledge, said to include collocations, grammatical functions, and associations among other components of word knowledge such as meaning, appropriateness and frequency. It is also useful to think of the depth of knowledge required to master these aspects of word knowledge in order to gain a more complete picture of the vocabulary knowledge required of learners. Regarding just one of these components, research shows that much language consists of sequences or chunks of words which operate as multi-word units, and as Moon (1997) argues, much of the input that we experience is comprised of formulaic sequences that cannot be understood simply by looking at the individual lexical items they are composed of.

Research into single vocabulary items that intermediate and advanced learners acquire has shown that they are learnt under conditions incidental to self-conscious learning (Brown, Waring, & Donkaewbua, 2008; Day, Omura, & Hiramatsu, 1991; Elley 1989; Horst, Cobb, & Meara, 1998; Jenkins, Stein, & Wysocki, 1984; Nagy, Herman, & Anderson, 1985; Pitts, White, & Krashen, 1989; Robbins & Ehri, 1994; Webb, 2007). That is to say, more or less unconscious learning of words is predominantly done through reading or listening. One of the exciting developments in recent years is that formulaic language or formulaic sequences, a long-standing topic of discussion by language teachers, is finally attracting the attention of researchers in applied linguistics and other disciplines (Alharthi, 2015; Meunier & Granger, 2008; Nessehauf, 2005; Wray, 2002). Emphasizing this important aspect of language knowledge, Schmitt (2004) edited an informative book that summarizes the research done in the last decade in the area of formulaic sequences, itself the focus of research in the field of phraseology.

A number of different terms for various types of multi-word units have been found in the empirical and theoretical literature but the most commonly used are lexical chunks and lexical phrases (Schmitt, 2000). Related to this, a variety of multi-word units in English have been researched with L2 learners including sub-types such as collocations (the tendency for words to occur together in discourse), binomials, idioms and lexical bundles (Conklin & Schmitt, 2008; Huang, 2015; Macis & Schmitt, 2016; Nguyen & Webb, 2016; Pellicer-Sánchez, 2015; Webb, Newton, & Chang, 2013). As one of the major formulaic variations and a corpus accessible aspect of phraseology, collocation is currently associated with improvements in fluency and the productive use of words (Boers, Eyckmans, Kappel, Stengers, & Demecheleer, 2006; Crossley, Salsbury, & McNamara, 2015; Schmitt, 2008). In other words, these scholars assert the necessity for an L2 learner who wants to use language accurately and fluently to have knowledge of collocation as one of the primary aspects of the use of a word since “knowing a word involves knowing what words it typically occurs with” (Nation, 2001, p. 56). Collocation has received much attention from researchers for both its linguistic and its pedagogical implications, and it is thus no surprise that González Fernández and Schmitt (2015) observe that “... perhaps the most studied category in applied linguistics is collocation” (p. 95).

According to Hill (2000), two thirds of collocations are produced in spoken and written discourse and the proportion of collocations in L2 input is higher than that of single-word items.

Despite the benefits for L2 learners of having knowledge of collocations since it would help them to achieve greater levels of accuracy and fluency in L2, previous research suggests that L2 learners' knowledge of collocations is less than their knowledge of single-word items (Bahns & Eldaw, 1993). Given L2 learners' lack of familiarity with collocations compared with single-word items, we would expect that the acquisition of collocations tends to be slow (Altenberg & Granger, 2001; Kuiper, Columbus, & Schmitt, 2009; Laufer & Waldman, 2011) and an area of particular difficulty for adult L2 learners and consequently an aspect of vocabulary knowledge in need of further empirical investigation.

For the past decade, we have been witnessing a heated debate over two main approaches of vocabulary learning and teaching, namely intentional (i.e. directing learners' attention to target words) and incidental (i.e. learners forming associations among words without consciously attending to learning). Regarding formulaic language, the majority of empirical studies has focused primarily on intentional approaches to the teaching of collocations and applying insights to the L2/EFL classroom (Boers, Demecheleer, Coxhead, & Webb, 2014; Chan & Liou, 2005; Jones & Haywood, 2004; Laufer & Girsai, 2008; Lindstromberg & Boers, 2008; Nguyen & Webb, 2016; Peters, 2014, 2015; Sun & Wang, 2003; Webb & Kagimoto, 2009). These studies have produced results in support of the explicit approach to promoting learners' knowledge of collocations. However, there remain the issues of teaching all the collocations that learners will need within the constraints of a timetabled course, in other words, which collocations are worth spending time on? And how should such learning be delivered in the classroom? Besides, one might argue that the input L2 learners have been most exposed to is the language of instruction, received in a classroom. This unconscious focus on (by teachers) or experience of (by learners) classroom language is reinforced by the fact that most EFL teachers are not aware of the importance of collocations and so the conscious focus of their teaching is on individual words (Garnier & Schmitt, 2016). That is, since the majority of English instructors are L2 speakers themselves, there is the likelihood that their students' exposure to collocations will be rather limited (Meunier, 2012). This might explain the lack of any significant effect on learners' knowledge of phrasal verbs regardless of the type of training and hours of classroom input they have received (Schmitt & Redwood, 2011). It is worth reiterating here that collocations are learned incrementally through multiple exposures and at different rates; see Alharthi (2014) for a review of the incremental nature of acquiring vocabulary knowledge and Henriksen (2013) for a review of the learning of collocations.

Put differently, collocations as one type of contextualized word knowledge are typically implicit and difficult to explain and are likely gained from massive exposure to the target language (Schmitt, 2014). Learners need therefore to adopt other incidental approaches and be prepared for the learning of collocations outside class. Due to recent advancements in information technology, students living in the twenty-first century have shown patterns of literacy development different from those living in the twentieth century. The internet has become a familiar tool of communication and learning for today's students. As a result, they tend to use online source materials more extensively and more frequently than before, rather than using traditional print media such as textbooks for learning English. This is especially true for electronic communication, nowadays widespread, and so EFL environments may no longer be as impoverished as they used to be. EFL learners are thus fortunate to be able to actively engage in learning outside the classroom

and set language books, and also access language input on TV, in music videos and various forms of social media, all of which in turn may help them with learning collocations.

It is worth noting that only a few empirical studies have been published reporting strong implications for incidental/implicit learning of collocation such as, for example, Webb, Newton and Chang (2013) who reported that the implementation of aural support during a reading task was an effective approach leading to significant learning gains of collocation knowledge. In a study of incidental learning of collocation, Durrant and Schmitt (2010) found that adult learners retained information about what words appear together in the input to which they were exposed. A study by Pellicer-Sánchez (2015) reported reading to be another type of implicit task to positively contribute to the learning of collocations. She found that the knowledge of orthography was learned incidentally through reading and at a rate similar to the form and meaning of individual words. While it can be safely assumed that these studies are useful in terms of raising learners' awareness of the incidental learning of collocation through reading and listening, it needs to be acknowledged that to date little empirical evidence has been produced to support the effectiveness of the incidental/implicit approach to learning collocations through reading. The goal of the classroom study we report in this paper was to help fill this gap in our knowledge. One important and relevant question that needs to be answered is what are the effective factors that may lead to the mastery of collocations? We shall therefore review the topics of frequency and out-of-class exposure to collocations in the following two sections.

## Review of research

### *The effect of frequency on L2 collocations learning*

A number of researchers have argued that multi-word units, i.e. collocations, formulaic sequences and lexical bundles, play a crucial role in both language use and language learning (Meunier & Granger, 2008; Nattinger & DeCarrico, 1992; Schmitt, 2004; Wood, 2010; Wray, 2002). Collocation is an integral element of vocabulary knowledge and lexis comprises the basis of any L2 English program; however, little research has looked beyond individual lexical items to explore the incidental learning of collocations in classroom-based settings. The identification of collocations is commonly explained through two approaches: the first, the phraseological approach, which is based on a free combination of semantically transparent words (e.g. Cowie, 1994; Howarth, 1998; Nesselhauf, 2005) for example, words such as *powerful* and *strong* have similar meanings, *engine* collocates with *powerful* but not with *strong* due to restricted co-occurrence; the second, the frequency approach, which is based on word combinations that co-occur more frequently than could be expected by chance. This approach is also known as statistical approach which involves various formulas to search corpora and identify the words that pattern together (e.g. McEnery & Wilson, 2001; Sinclair, 2004). For example, words such as *support*, *experience*, *confidence*, *control* all share mutual information scores with *gain*, indicating their co-occurrence, therefore *gain support*, *gain experience*, *gain confidence* and *gain control* are all collocations. This latter approach might be helpful for researchers, teachers and learners to use in research as it "requires the most objective measure to record the frequency of collocations" (Nguyen & Webb, 2016, p. 3) and is thus the type of approach that we will adopt in the present study. On the other hand, a shortcoming of the frequency-based approach is that it does not include other semantic factors such as concreteness of meaning (Walker & Hulme, 1999), congruency (Wolter & Gyllstad, 2011, 2013; Yamashita & Jiang, 2010), and transparency (Moon, 1998;

Nesselhauf, 2003) that may affect the ease with which the collocations may be learned. This approach is also beneficial as any subjectivity in identifying the target collocations is eliminated. Since frequency of occurrence has been widely accepted as one of the valid predictors of learning single-item vocabulary (Nation, 2001; Schmitt, 2010), it is likely that the same holds true for formulaic sequences such as collocations. For example, Peters (2014) found that there was a remarkable repetition effect with 1, 3 and 5 exposures and this led to improved learning of target collocations. These results appear to be consistent with the findings of Durrant and Schmitt's (2009) study where adult non-natives were exposed to target collocations of different levels of frequency from the British National Corpus (BNC), whose authors concluded that learners tended to acquire more frequent collocations before less frequent ones. In a similar vein, Nguyen and Webb (2016) investigated the relationship between the learning of verb-noun and adjective-noun combinations and frequency by Vietnamese undergraduate EFL learners at three-word frequency levels. They demonstrated a clear trend in that the collocations with the lowest frequency were learned less well.

The discussion above suggests that the frequency of collocations in a given language is likely to be an important factor in the learning of collocations. The present study seeks to shed some light on frequency as a potential factor in L2 learners' knowledge of collocations.

### ***The effect of out-of-class exposure on L2 collocations learning***

Another trend we can discern in the studies of L2 collocations is that learners' engagement with the target language seems to be a powerful facilitator for learning collocations. Indeed, research has shown that engagement activities such as reading, watching TV and social networking in English best enhance the effectiveness of L2 speakers learning English collocations. However, it is surprising how few empirical studies have so far been conducted to examine this particular variable.

One of the few attempts to do this is Schmitt and Redwood's (2011) study who found that certain out-of-class exposure to English, for example via extensive reading and watching English language films and TV programs, facilitated the learning of phrasal verbs. Along the same lines, González Fernández and Schmitt (2015) conducted one of the most thorough investigations of the acquisition of collocations in L2 contexts. Their results showed that activities that provide input in natural forms such as reading texts, watching TV, listening to music and using social media appeared to have a positive effect on the acquisition of collocations. Quite recently, Garnier and Schmitt (2016) made a similar case for the learning of polysemous phrasal verbs in English by 128 Chilean EFL learners. The effect of L2 engagement in leisure activities on the learning of phrasal verbs is clearly apparent from their results. Specifically, Garnier and Schmitt (2016) summarized their findings by saying "both reading and social networking seemed to promote the learning of PVs. This is good news as it suggests that it is possible to learn a lot outside the classroom, via daily activities that are engaging and enjoyable" (p. 38). While these three studies suggest that there might be instances where the L2 learners face formulaic sequences such as collocations which they have not encountered in the classroom, other sources such as watching TV programs, listening to music or radio and using social media may have significantly contributed to better knowledge of L2 speakers' collocations. In fact, such out-of-class activities have gained in popularity among our students. Because formulaic sequences like collocations commonly occur in

informal input, these out-of-class activities are believed to be particularly appropriate for learning collocations. However, to the best of our knowledge few studies, with the exception of the above mentioned ones, have addressed the question whether out-of-class activities such as watching TV, listening to radio or music, reading English books and using social media play a significant role in the learning of target collocations. In order to tackle this issue, the current study explored how much effect these kinds of activities have on the learning of collocations.

### **Aims and research questions**

The present study seeks to contribute to knowledge about the effectiveness of reading for the incidental learning of collocations by non-native English speakers. We will thus establish the relationship between the frequency of target collocations and the degree to which EFL learners know them. Basing our research on Pellicer-Sánchez's (2015) study intended to assess collocations gains; we will include pseudo-words rather than a pre-test/post-test design. To obtain further insight into the issue of incidental learning of collocations, we will also seek to establish the effect of some other learner-internal factors on such learning. In light of the studies reported in the literature, we posed the following research questions:

1. How does the element of frequency in reading relate to the EFL learners' incidental knowledge of collocations?
2. How does incidental learning through personal language use outside the classroom relate to the EFL learners' knowledge of collocations?

### **Context and methodology**

#### ***Participants***

The participants in this exploratory study of incidental collocation learning were 46 Arabic-speaking young adult EFL learners recruited from two intact sophomore English classes at King Abdulaziz University (KAU), Saudi Arabia. The participants' average age was 20 and they had two (three-hour) compulsory courses in their first year BA program focusing on reading. They share a similar educational background, that is, all study participants had considerable language learning experience since they had been learning English for at least ten years prior to the present study being carried out. At university, they learned EFL through the medium of classroom instruction, with the approach having been a communicative one but with a special emphasis on grammar. Prior to commencing their degree at KAU, they achieved the necessary minimum scores in the university entrance English examination. This test is similar to a standardized test called the Michigan Test where three main language components are measured, namely vocabulary, grammar and writing. The study participants' overall level of competence was estimated to be between intermediate and upper-intermediate which corresponds to the level between B1 and B2 of the Common European Framework of Reference (CEFR).

As the main purpose of the present research is to investigate the knowledge of collocation of EFL learners, it was necessary to choose participants with advanced proficiency levels in English. Specifically, to gain an accurate estimate of the study participants' vocabulary proficiency and to ensure they have the ability to complete the research test, they sat Vocabulary Size Test (VST) (Nation & Beglar, 2007) and obtained a receptive vocabulary size of 5,000+ word families which

enables them to understand the running words in the test. Consequently, this brief account describes a typical language learning background for Saudi learners of English of this age cohort.

### *Selecting the target collocations*

As the present study focuses on two-word collocation combinations, we decided to limit the selection to verb-noun combinations. Specifically verb-noun collocations were chosen as they are likely to be more problematic than other types of lexical collocations (Chan & Liou, 2005; Nesselhauf, 2003; Yamashita & Jiang, 2010) as well as the most frequently researched collocation type (Paquot & Granger, 2012). The initial step was to consult a range of different sources which listed the target collocations as there is no established and comprehensive list of collocations in existence to base our research on (Macis & Schmitt, 2016). Moreover, we wished to focus on high-frequency candidate items as we would like to know how familiar the study participants were with the type of collocations they would presumably have had the most exposure to. Over 44 potential target collocations were provisionally selected and included in a candidate pool from three passages in a textbook intended to be taught in the study participants' BA course called *Pathways 3: Reading, Writing, & Critical Thinking* (Vargo & Blass, 2014) published by National Geographic Society. We focused on the collocations that were embedded in texts without being highlighted and clickable, so that there is a possibility of learning them incidentally.

After this process, a rigorous set of criteria was adopted in selecting the final list of collocations. To avoid the effect of expected selection bias, the selected collocations were then verified by checking them against collocation-specialized dictionaries, namely the *Oxford collocations dictionary for students of English* (Lea, Crowther, & Dignen, 2002) and the *Longman collocations dictionary and thesaurus* (2013). Following the procedure established by Peters (2015), only 21 verb-noun combinations that occurred in one of the above mentioned dictionaries were accepted as verified collocations, and hence included in the present study. Next, the *Corpus of Contemporary American English* (COCA, Davies, 2008) website at <http://www.corpus.buy.edu/> was consulted as a reference corpus, and it was found that all 21 collocations were listed in the COCA database. To increase the level of ecological validity, two raters, who were teaching English at the time and who were also unaware of the aim of the study, were asked to confirm that the collocations selected had not been previously taught. Only lexical collocations on which the two raters agreed as unknown by students were maintained. With regard to the optimal level of linguistic difficulty, the target collocations comprised collocations of the same or higher frequency levels within the most frequent 1,000 word families, the second most frequent 2,000 word families and the third most frequent 3,000 words families from Nation's (2012) BNC website (<http://www.victoria.ac.nz/lals/about/staff/paul-nation>). We chose this new list as it represents everyday language, is regularly updated, and comprises lexical items used nowadays by native speakers than outdated lists such as the General Service List (West, 1953). For example, both the node word *friend* and the collocate *make* in the collocation *make friend*, are listed in the 1,000 most frequent band. On the other hand, in the collocation *share source*, the node word *source* is drawn from the 3,000 most frequent band and the collocate *share* is from the 1,000 most frequent band.

***Apparatus: Collocation test***

The primary goal of the present study was to establish the learners' knowledge of the target collocations, and a meaning-recall cloze test was designed to fulfill this purpose. The instrument consisted of gapped sentences where 21 potential target verb-noun collocations were inserted into an off-line pencil-and-paper test. A meaning-recall cloze format was used in the present study due to the lack of guessing effects compared with a multiple-choice recognition format. One of the possibilities was to create a test assessing the learners' knowledge of collocations using a translation production test similar to the one adopted in Webb, Newton, and Chang's (2013) study. It could be argued that measuring the acquisition of collocation knowledge using a translation production format is typically a more demanding task and may be highly dependent on the knowledge, expertise and accuracy of the learner to identify and spot answers from L2 to L1 and vice versa. As a result, this type of measurement would give a misleading picture of collocation learning. Conversely, the recall cloze format has been proven to be an effective instrument as used in recent studies (e.g. Garnier & Schmitt, 2016; González Fernández & Schmitt, 2015) and ultimately the creation of another measure was beyond the scope of the current study. The study participants were provided with a target collocation embedded in a sentence of English (the participants' FL). In this sentence, the context was left intact but contained two gaps which corresponded to each of the two content words which formed the collocation measured. The test format does not require the test-takers to choose from a list of words as in the case of an insert-the-word exercise where a list of target words is given. Therefore, to constrain the choice of the potential collocations in each blank, the first letters of the two words were supplied and marked in bold. The verbs to be completed were all in the infinitive, so participants were not restricted to using inflectional morphology. The meaning of the target collocations can be looked up with the contextually appropriate denotative meaning presented between brackets and printed underlined. The following examples illustrate the test format:

Humans tend to live with each other and **a**\_\_\_\_\_ **c**\_\_\_\_\_. (Keep out from disagreement)  
 It was easy for him to **m**\_\_\_\_\_ **f**\_\_\_\_\_ with people from all walks of life. (Begin a friendship with someone)

The extended surrounding contexts were submitted to frequency analysis through the *The Compleat Lexical Tutor* (Cobb, 2017), to ensure that test-takers would not have problems reading and understanding the running words.

***Reading materials***

A methodological concern was how to include the target collocations in a rigorous test to assess the study participants' pre-knowledge while ensuring that they would be unknown to them. In order to do so, prior knowledge of the collocations was controlled for by means of non-words (sometimes called pseudo-words). This is an effective way of controlling for prior knowledge of target invented items in advance and of ascertaining that pseudo-words are comparable to real words. This is designed to not alert study participants to any previous knowledge of the target items. We decided to include only 10 pseudo-words in a modified text to create reading materials that closely resembled authentic ones where we changed one or more letters in the second component of the target verb-noun collocations (e.g. *made-descuvory\**). This was carefully

constructed to ensure the study participants would not give up indicating whether the collocate was a real English word or not. The text of the reading task was obtained from a website called Voice of America. It is a two-page narrative text modified to comprise ten pseudo-word collocations. The text narrates a fictional story that takes place in the United States and falls into the category of texts within the 2,000 most frequent words. This was considered crucial for the study participants to have a chance of comprehending it thoroughly. The pseudo-words were added without too much altering the style of the text, its nature and general structure and leaving it sufficiently spaced out so as not to be too prominent. That is, the pseudo-words were carefully distributed throughout the text which was then revised by a native speaker to ensure that it still conveyed the same content and its style was not distorted.

### ***Biodata survey***

A self-report survey was the second main data gathering instrument used in the present study. It aimed at finding out personal information about the study participants' language learning experience and more specifically their amount of language exposure through incidental learning. The survey was adapted from Garnier and Schmitt's (2016) questionnaire and consisted of 10 closed items. As McDonough and McDonough (1997) point out, "It is useful for the majority of the questions to be answered by ticking a box or circling an alternative to enable easier counting" (p. 174). Therefore, a 10-item lie scale was also implemented in order to check to what extent the study participants showed the amount and type of exposure they had had to English in their responses. The 5-point scale used for rating the participants' frequency of using English outside the classroom was as follows: 1= less than once a month, 2= several times a month, 3= once a week, 4= several times a week, 5= every day. It was content validated as it was professionally constructed by the second author, Norbert Schmitt, who is an expert in designing such a study. Well validated surveys are often considered acceptable without further checking of reliability, which is a bit difficult to do since it requires the instruments to be administered twice to the same participants with a gap of time. The first section includes items to gather demographic information such as the study participants' age, number of years of schooling in which they had studied English as a FL and self-assessed proficiency in English. The second section comprised the remaining set of items which focused on opportunities to use English either socially or alone, therefore activities like extensive reading, watching English videos or TV, listening to English music or radio stations and using social networking were considered in this situation.

### ***Procedure and data analysis***

The current study was conducted in four sessions on KAU premises under the supervision of the present researcher over a period of three weeks (see Figure 1 for a schematic description of the study procedure).

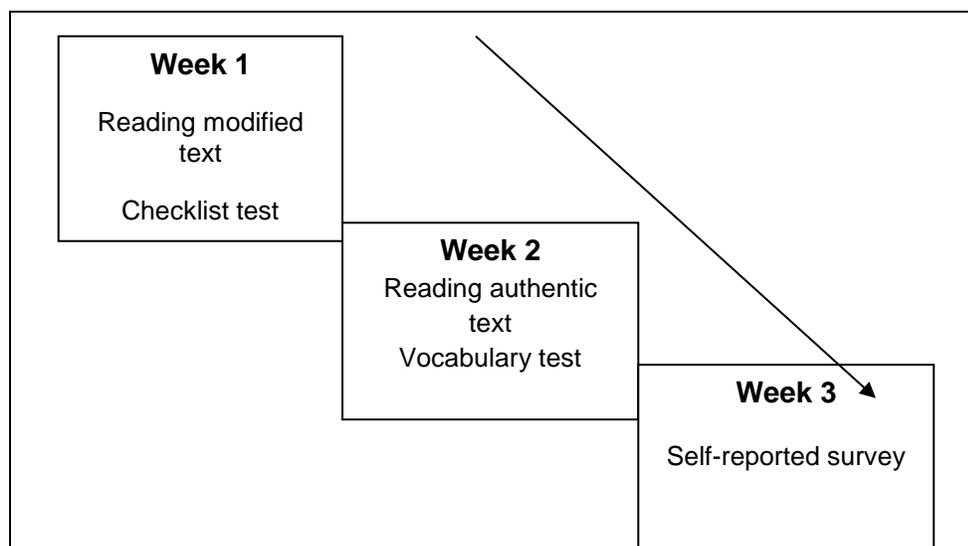


Figure 1. Schematic description of the study procedure

Prior to the beginning of the sessions the study participants were informed about the nature of the study, and that the information gathered would be treated confidentially. In the first week, the participants were assigned a reading task during normal class time and were instructed to read a modified text containing 10 pseudo-word collocations for general comprehension. After reading the text, a checklist test with the potential pseudo-word collocations was administered to the study participants to confirm that they really did not know the target collocations. That is, the study participants were asked to mark the collocations that they identified as real English collocations and leave those alone which they believed were not real. The checklist tests were scored as follows: 0= incorrect and 1= correct. The outcomes showed virtually no prior knowledge and no significant differences in the study participants' scores. A week later during session two, the study participants were provided with the authentic reading text that contained the real 21 target collocations and were informed to read it intensively to be ready to answer comprehension questions afterwards; however, there was no mention of the vocabulary test. After a 15 minute break, vocabulary tests were then unexpectedly administered to the study participants in the third session which took an hour to complete. The same procedure was adopted in scoring the vocabulary test: each item was assigned one point and missing or incorrect answers received zero, totaling 21 points for the whole test. However, accurate spelling was not required for a collocation to be marked as correct as long as the meaning definition was clear and understandable. To avoid a possible fatigue effect, the study participants were given the surveys in the final week, with the session lasting approximately 15-20 minutes.

## Results

We are able to respond to research question one by using the descriptive statistics of the study participants' performance on the meaning-recall cloze test in which they were required to supply the missing collocations (see Table 1). On closer inspection of Table 1, we can see that the study participants were able to produce only a modest learning gain, with the highest mean score gain obtained at the first frequency level with 14.68 out of 21 collocations (92.18%); 12.55 out of 21 collocations (85.43%) at the second frequency level and 9.75 out of 21 collocations (69.70%) at

the third frequency level; see below for further discussion. The results of a one-way ANOVA showed statistically significant differences in collocation learning ( $F = 32.94, p = .001$ ), although it represents only a very small increase which suggests that frequency of occurrence contributes very little to the successful learning of collocations

Table 1

*Descriptive statistics of knowledge of collocations*

Test scores (21 items)	1K	2K	3K
Mean	14.68	12.55	9.75
Standard Deviation	2.86	3.13	2.71
Minimum	9	7	4
Maximum	18	13	8

*Note.* The total number of tested items at 1k, 2k, and 3k was 7. 1k= 1000 level, 2k= 2000 level, 3k= 3000 level.

The relatively small standard deviations also point to the small differences between the study participants' scores at each collocation level. This is shown in Figure 2, which displays the distribution of the collocation test scores. Most study participants scored between 9 and 16 but none scored the full mark, while two individuals scored 0, which contributes to showing how hard it may be for these EFL learners to learn the meanings of verb-noun collocations.

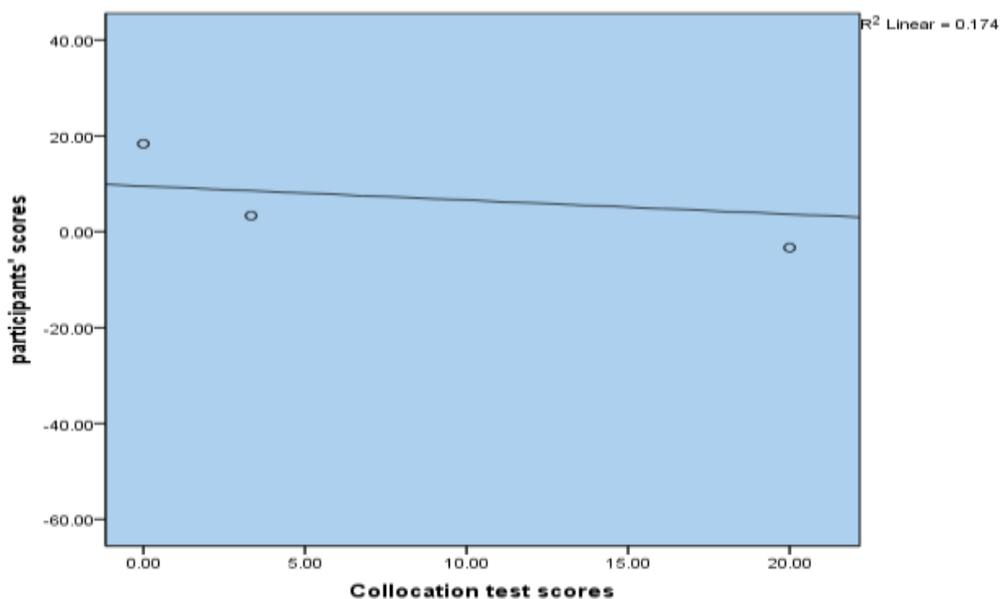


Figure 2. The distribution of collocation test scores

In absolute terms, the increase in the command of collocations appears very slow and drops consistently from one thousand to the next as one would expect. It seems that higher frequencies indicate higher levels of exposure for the study participants. In contrast, they know less than half

of the portions at the third frequency level which strongly suggests that they know few collocations at this level, and consequently may have a serious comprehension deficit. Overall, it appears that the study participants' knowledge of collocations is dependent on the frequency of a given collocation and that the collocations with the highest frequency are likely to be learned rather than the more infrequent ones.

This raises the question of what other factors, e.g. out-of-class exposures such as watching TV programs, listening to radio or music, reading English books and using social media, beside frequency, contribute to the learning of collocations. To answer this question, a Multiple Regression (MR) analysis was carried out to examine whether the above variables in combination could predict the study participants' success in learning collocations. The MR analysis was performed with each of the predictor variables entered into the regression equation in a different order. That is, all of the orders of entry of the four predictors were performed and the effect of each variable was compared to all others.

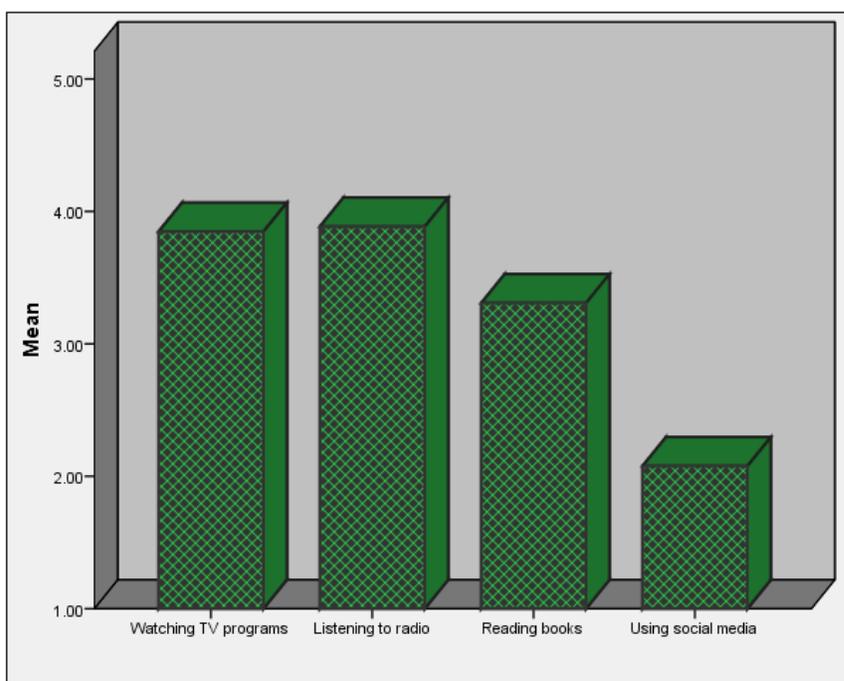


Figure 3. Mean scores for out-of-class activities in the incidental learning of collocations

Figure 3 displays the mean scores for out-of-class activities in the incidental learning of collocations by the study participants. The findings of the self-reported survey showed that involvement with collocations in everyday communication in the target language may well be just as essential a factor as frequency of occurrence of given collocations. That is, the results indicate that the study participants seemed to be motivated to find different input activities such as listening to radio and watching TV and reading books which fall between the frequencies of “once a week” and “several times a week”. By contrast, using social media received a low frequency rating of “several times a month”.

Table 2

*Predictor variables of correct scores in collocation test*

Predictor variable	Dependent variable	Beta	Sig
Watching TV programs	correct score in collocation test	.276	<.001
Reading English books	correct score in collocation test	.488	<.001

The results of the MR analyses (see Table 2) indicate that in building regression models with any four predictors, regardless of the combinations and what the orders of entry, it turned out that watching TV programs, listening to radio and reading English books together account for 53% of the variance in the collocation test performance. These findings revealed statistically significant positive relationships between watching TV programs (Beta .276,  $p < .001$ ), listening to radio (Beta .343,  $p < .001$ ), reading English books (Beta .488,  $p < .001$ ) and participants' test scores and indicate a positive contribution to collocation knowledge. This means that the more the study participants encounter collocations in such everyday communicative situations (watching TV, listening to radio and reading books), the more they are successful in maximizing their repertoire of collocations. However, there was a non-significant relationship between using social media and participants' knowledge of collocations.

## Discussion

A review of the current literature on L2 single-word items indicates that reading fosters both vocabulary knowledge (e.g. Horst et al., 1998) and language proficiency (e.g. Renandya et al., 1999). The findings of this study have demonstrated that although EFL undergraduate learners showed limited knowledge of L2 collocations, reading became an effective facilitator of the incidental learning of verb-noun collocations. This supports the results of studies by Durrant and Schmitt (2010), Pellicer-Sánchez (2015) and Webb, Newton and Chang (2013). The findings of the current study have also pointed to a general trend of frequency of occurrence of given collocations leading to some extent to the learning of collocations. That is, higher frequency verb-noun collocations were inevitably learned more readily by most participants than lower frequency verb-noun collocations. These results extend existing reading research to the field of incidental collocation learning by L2 learners by showing that frequency of occurrence is indeed an adequate predictor of the uptake of formulaic sequences and did account for differences in the growth of the knowledge of collocations while reading (Pellicer-Sánchez, 2015; Garnier & Schmitt, 2016; Schmitt & Redwood, 2011). Compared to the studies by Webb, Newton, and Chang (2013) and Pellicer-Sánchez (2015), where retention periods were measured in terms of pre-test/post-tests and over longer periods, the present study has the advantage of mastery of a verb-noun collocation inventory even though it lacked sufficiently precise pre-test measures when the post-test was administered one week after the pseudo-words checklist session.

A further point to keep in mind is that the study participants had insufficient knowledge of instances of collocation. It will be recalled from the literature review that growth in the knowledge of collocations tends to be slow and that L2 learners encounter a considerably lower frequency of occurrence of collocations in their L2 input (Altenberg & Granger, 2001; Columbus & Schmitt,

2009; Howarth, 1998; Kuiper, Li, & Schmitt, 2010; Laufer & Waldman, 2011). This would be congruent with the argument espoused by Webb, Newton, and Chang (2013) that the potential for learning the majority of target collocations incidentally is limited. The findings in the present study also provide us with the opportunity to observe how well the participants learn new verb-noun collocations in real terms. It is important to note that the success with which the study participants identified new collocations can be compared with Nguyen and Webb's (2016) study because the settings of the two studies are similar (EFL) as is the investigation of the mastery of collocations at the first three 1,000 word frequency levels. Although the large number of tested items and the relatively easy test format of target collocations explored in Nguyen and Webb's (2016) study were different from those of the present study, the mean scores of learning gains across the three levels of verb-noun combinations obtained in Nguyen and Webb's (2016) study were to some extent similar to those of the present study (see Table 3).

Table 3

*Learning gains in collocations, compared with Nguyen & Webb's (2016) study*

<b>Study</b>	<b>1K</b>	<b>2K</b>	<b>3K</b>
Nguyen & Webb (2016)	16.80	13.93	10.42
Current study	14.68	12.55	9.75

The mean scores of the current study show the same trend as Nguyen and Webb's (2016) study as they decrease successively over the three word frequency levels. However, the scores obtained in Nguyen and Webb's study were slightly higher than those obtained in the present study, given the assessment measures employed in their study were designed to measure the participants' receptive knowledge of collocations. In contrast, the testing format of the present study is open response and requires considerable knowledge of collocations. That is, participants must actively recall and produce the missing items in a given context, which in turn could have increased the difficulty of the task and have yielded even lower scores. In other words, having the 21 collocations embedded in sentences and L2 learners being unaware of the unusual combinations of these words made their task rather difficult and caused comprehension problems for them. Additionally, findings of recent collocation research use the more demanding productive tests usually reveal good knowledge of collocations, for example, 40% by Garnier and Schmitt (2016), 26% by Webb, Newton, and Chang (2013), 48% by Schmitt and Redwood (2011), and 33% by Macis and Schmitt (2016). However, a word of caution is needed in the interpretation of such studies. The conflicting findings of these studies may be attributed to differences in participants (e.g. different language competence and different L1 background), different types of formulaic sequences and different condition underlying the tasks set. Compared to these figures, the results of the current study are less encouraging and not so impressive, given that the study participants are English majors and could be expected to show similar trends in their knowledge of collocations as the college students studying for their BA in the above studies.

The conflicting results might be explained with reference to Granger's (1998) finding that learners tend to focus on a small number of safe formulaic consequences that they felt confident producing. As Peters (2015, p. 3) notes, "... FL learners might not allocate sufficient attentional resources to the verb when they encounter verb-noun collocations." That is, the study participants might have been familiar with the verb-noun combination *organize a party*, but have to draw on

limited knowledge when asked to supply the correct combination *crash a party*. As collocations, e.g. verb-noun combinations, are essential blocks for language use, but unfortunately problematic for learners. Thus Saudi EFL learners need to know that verbs can have different meanings when they are combined with particular nouns. One might argue that formulaic sequences, at least verb-noun collocations, tend to be less transparent and salient in language input if learners are habituated to define and process them in the form of their constituent words (Wray, 2002). There seems to be a particular morphological feature that distinguishes verb-noun collocations from adjective-noun or adverb-adjectives combinations. Along this line, Laufer (2011) found that EFL learners struggle to locate verb-noun collocations in dictionaries as they occur in a variety of inflections. She provided an example of a verb-noun collocation (*to take measures*) in two different forms: as *they took strong measures against dangerous drivers* where the collocation occurs with the verb in the past, and as *measures are being taken to reduce crime in the city* where the verb appears in the passive. This in turn shows that variation in morphology is often seen as a factor affecting the ability of learners to recognize or interpret a collocation. It is not unexpected that language input in FL classrooms such as in Saudi Arabia is presented to the learners by textbooks and teachers. Although there is no reason to think that the target verb-noun collocations in the current study would benefit from an explicit approach by the study participants, at least not in any systematic way, most talk in FL environments is delivered by non-native teachers. These may lack awareness of the importance of collocations and may therefore not provide a wider range of collocations in terms of their frequency of occurrence and therefore fail to help learners expand their knowledge of collocations.

The current study has also explored the influence of exposure related factors on the participants' productive knowledge of verb-noun collocations. And yet, according to the MR analysis, watching TV programs, listening to radio and reading English books did have significant effects on the acquisition of verb-noun collocations. These findings are echoed by other research such as by Garnier and Schmitt (2016), González Fernández and Schmitt (2015), Macis and Schmitt (2016), and Schmitt and Redwood (2011). Unsurprisingly, EFL learners can be exposed to different sources of English with a potentially wide range of vocabulary including collocations needed for communicating in English. The study findings reflect the fact that the participants were aware of the importance of encountering different types of formulaic sequence such as collocations outside the classroom through incidental learning. Furthermore, research has shown that L2 learners are likely to take advantage of opportunities to interact and socialize with native speakers to practice their language which in turn has an effect on their knowledge of collocations (Garnier & Schmitt, 2016; González Fernández & Schmitt, 2015). Interestingly, the findings of the present study did not reveal that social media had a positive impact on the productive knowledge of collocations. Perhaps language input such as that provided by social media would be clearly correlated with EFL learners' knowledge of verb-noun collocations if the latter were measured in a receptive / recognition test of collocations.

## Conclusion

This paper investigated the degree of sophomore English majors' productive knowledge of verb-noun collocations at three word frequency levels while reading. It also aimed at exploring the relationship between out-of-class exposure-based variables and the learning of verb-noun collocations. The findings suggest that collocations can be learned incidentally through reading

and that high frequency verb-noun collocations tend to be easier to learn than low-frequency collocations. Interestingly, the development of the knowledge of collocations is effected by the frequency of occurrence of collocations in a given context. The results from the biodata survey showed that watching TV programs, listening to radio and reading English books was positively correlated with the ability to predict learning gains in the productive knowledge of collocations.

As is the case with any study, there are a number of shortcomings that need to be noted in respect of the current study. These shortcomings are presented so that other researchers will be able to take them into consideration whenever they carry out research in the same field. Although the present study did not incorporate a pre-test in order to direct participants' awareness to target collocations, there may be a validity concern about the use of pseudo-words. That is, it may lead some test-takers to identify pseudo-words as real words which would increase their false alarm rate. To ensure that target items are unknown to all participants, we suggest that future researchers assign and administer the test to a cohort of participants that represent the individuals who take part in the real study. Also, the target items included in the collocation test were limited to only verb-noun combinations and administered to only 21 candidates. Future studies might include more sample items and investigate the extent of learners' knowledge of other types of collocation to reach more robust conclusions. The current study tested the productive knowledge of collocations and did not test receptive knowledge. Further studies might measure the knowledge of collocations both receptively and productively as the former resembles the receptive skills needed in reading and listening. In fact, previous research has shown that receptive knowledge of formulaic language seems to be a powerful facilitator of fluency in reading (e.g. Conklin & Schmitt, 2008; Ellis, Simpson-Vlach, & Maynard, 2008).

#### About the Author:

**Thamer Alharthi** is Associate Professor of Applied Linguistics at King Abdulaziz University, Jeddah, Saudi Arabia. His research interests, originating from the years he spent studying at Essex University, are in the areas of vocabulary attrition, acquisition and teaching, with a specific focus on vocabulary development. He has also worked on projects exploring the role of formulaic sequences in foreign language use. **ORCID.org/0000-0002-6914-418X**

#### References

- Alharthi, T. (2014). The Dynamic and Incremental Features of Vocabulary Acquisition. *International Journal of English Linguistics*, 4, 70–77.
- Alharthi, T. (2015). Adding More Fuel to the Fire: A study of attrition in formulaic sequences by adult learners. *Arab World English Journal*, 6, (3), 230–243.
- Altenberg, B., & Granger, S. (2001). The Grammatical and Lexical Patterning of MAKE in Native and Non-native Student Writing. *Applied Linguistics*, 22, 173–195.
- Bahns, J., & Eldaw, M. (1993). Should We Teach EFL Students Collocations? *System*, 21, 101–114.
- Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the Effects of Exercise on Verb-noun Collocations. *Language Teaching Research*, 18, 54–74.
- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic Sequences and Perceived Oral Proficiency: Putting a lexical approach to the test. *Language Teaching Research*, 10, 245–261.

- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental Vocabulary Acquisition from Reading, Reading-while-listening, and Listening to Stories. *Reading in a Foreign Language*, 20, 136–163.
- Chan, T. P., & Liou, H. C. (2005). Effects of Web-based Concordancing Instruction on EFL Students' Learning of Verb-noun Collocations. *Computer Assisted Language Learning*, 18, 231–251.
- Conklin, K., & Schmitt, N. (2008). Formulaic Sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, 29, 72–89.
- Cowie, A. P. (Ed.). (1994). *Phraseology*. Oxford: Pergamon.
- Crossley, A. S., Salsbury, T., & McNamara, D. S. (2015). Assessing Lexical Proficiency Using Analytic Ratings: A case for collocation accuracy. *Applied Linguistics*, 36, 570–590.
- Davies, M. (2008). The Corpus of Contemporary American English: 450 million words, 1990–present. Retrieved September 18, 2017 from <http://corpus.byu.edu/coca>
- Day, R., Omura, C., & Hiramatsu, M. (1991). Incidental EFL Vocabulary Learning and Reading. *Reading in a Foreign Language*, 7, (2), 541–551.
- Durrant, P., & Schmitt, N. (2010). Adult Learners' Retention of Collocations from Exposure. *Second Language Research*, 26, 163–188.
- Elley, W. B. (1989). Vocabulary Acquisition from Listening to Stories. *Reading Research Quarterly*, 24, 174–187.
- Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic Language in Native and Second Language Speakers: Psycholinguistics, corpus linguistics and TESOL. *TESOL Quarterly*, 42, 375–396.
- 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.
- Granger, S. (1998). Prefabricated Patterns in Advanced EFL Writing: Collocations and formulae. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 79–100). Oxford: Oxford University Press.
- Henriksen, B. (2013). Research on L2 Learners' Collocational Competence and Development: A progress report. In C. Bardel, C. Lindqvist & B. Laufer (Eds.), *L2 Vocabulary Acquisition, Knowledge, and Use* (pp. 29–56). Euro SLA Monograph Series 2. European Second Language Association (EuroSLA).
- Hill, J. (2000). Revising Priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching Collocation: Further development in the lexical approach* (pp. 47–69). Oxford: Oxford University Press.
- 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. (1998). Phraseology and Second Language Proficiency. *Applied Linguistics*, 19, 24–44.
- Huang, K. (2015). More Does Not Mean Better: Frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing. *System*, 53, 13e23. Available at <http://dx.doi.org/10.1016/j.system.2015.06.011>
- Jenkins, J. R., Stein, M. L., & Wysocki, K. (1984). Learning Vocabulary through Reading. *American Educational Research Journal*, 21, 767–787.

- Jones, M., & Haywood, S. (2004). Facilitating the Acquisition of Formulaic Sequences: An exploratory study. In N. Schmitt (Ed.), *Formulaic Sequences* (pp. 269–300). Amsterdam: John Benjamins.
- Kuiper, K., Columbus, G., & Schmitt, N. (2009). Acquiring Phrasal Vocabulary. In S. Foster-Cohen (Ed.), *Advances in Language Acquisition* (pp. 216–240). Basingstoke: Palgrave Macmillan.
- Laufer, B. (2011). The Contribution of Dictionary Use to the Production and Retention of Collocations in a Second Language. *International Journal of Lexicography*, 24, 29–49.
- Laufer, B., & Girsai, N. (2008). Form-focused Instruction in Second Language Vocabulary Learning: A case for contrastive analysis and translation. *Applied Linguistics*, 29, 1–23. doi:10.1093/applin/amn018
- Laufer, B., & Waldman, T. (2011). Verb-noun Collocations in Second-language Writing: A corpus analysis of learners' English. *Language Learning*, 61, 647–672. doi:10.1111/j.1467-9922.2010.00621.x
- Lea, D., Crowther, J., & Dignen, S. (Eds.). (2002). *Oxford Collocations Dictionary for Students of English*. Oxford: Oxford University Press.
- Lindstromberg, S., & Boers, F. (2008). The Mnemonic Effect of Noticing Alliteration in Lexical Chunks. *Applied Linguistics*, 29, 200–222.
- Longman Collocations Dictionary and Thesaurus*. (2013). Harlow: Pearson Education.
- Macis, M. & Schmitt, N. (2016). Not Just 'Small Potatoes': Knowledge of the idiomatic meanings of collocations. *Language Teaching Research*. 21. (3), 321–340. doi:10.1177/1362168816645957
- McDonough, J., & McDonough, S. H. (1997). *Research Methods for English Language Teachers*. London: Arnold.
- McEnery, T., & Wilson, A. (2001). *Corpus Linguistics* (2nd ed.). Edinburgh: Edinburgh University Press.
- Meunier, F. (2012). Formulaic Language and Language Teaching. *Annual Review of Applied Linguistics*, 32, 111–129. doi:10.1017/S0267190512000128
- Meunier, F., & Granger, S. (Eds.). (2008). *Phraseology in Foreign Language and Teaching*. Amsterdam: John Benjamins.
- Moon, R. (1997). Vocabulary Connections: Multi-word items in English. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 40–63). Cambridge: Cambridge University Press.
- Moon, R. (1998). *Fixed Expressions and Idioms in English*. Oxford: Oxford University Press.
- Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning Words from Context. *Reading Research Quarterly*, 20, 233–253.
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Boston, Mass: Heinle & Heinle.
- Nation, I. S. P. (2001). *Vocabulary Learning in Another Language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2012). *The BNC/COCA Word Family Lists*. Retrieved September 20, 2017 from <http://www.victoria.ac.nz/lals/about/staff/paul-nation>
- Nation, I. S. P., & Beglar, D. (2007). A Vocabulary Size Test. *The Language Teacher*, 31, (7), 9–13.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.

- Nesselhauf, N. (2003). The Use of Collocations by Advanced Learners of English and Some Implications for Teaching. *Applied Linguistics*, 24, 223–242.
- Nessehauf, N. (2005). *Collocations in a Learner Corpus*. Amsterdam: John Benjamins.
- Nguyen, T. & Webb, S. (2016). Examining Second Language Receptive Knowledge of Collocation and Factors that Affect Learning. *Language Teaching Research*, 21, (3), 298–320. doi:10.1177/1362168816639619
- Paquot, M., & Granger, S. (2012). Formulaic Language in Learner Corpora. *Annual Review of Applied Linguistics*, 32, 130–149.
- Pellicer-Sánchez, A. (2015). Learning L2 Collocations Incidentally from Reading. *Language Teaching Research*, 21, (3), 381–402. doi:10.1177/1362168815618428
- Peters, E. (2014). The Effects of Repetition and Time of Post-test Administration on EFL Learners' Form Recall of Single Words and Collocations. *Language Teaching Research*, 18, 75–94.
- Peters, E. (2015). The Learning Burden of Collocations: The role of interlexical and intralexical factors. *Language Teaching Research*, 20, 113–138.
- Pitts, M., White, H., & Krashen, S. D. (1989). Acquiring Second Language Vocabulary through Reading: A replication of the *Clockwork Orange* study using second language acquirers. *Reading in a Foreign Language*, 5, 271–276.
- 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.
- Robbins, C., & Ehri, L. C. (1994). Reading Storybooks to Kindergartners Helps Them Learn New Vocabulary Words. *Journal of Educational Psychology*, 86, (1), 54–64.
- Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Schmitt, N. (Ed.). (2004). *Formulaic Sequences: Acquisition, processing and use*. Amsterdam: John Benjamins. doi:10.1075/llt.9
- Schmitt, N. (2008). Instructed Second Language Vocabulary Learning. *Language Teaching Research*, 12, 329–363.
- Schmitt, N. (2010). *Researching Vocabulary: A vocabulary research manual*. Basingstoke: Palgrave Macmillan.
- Schmitt, N. (2014). Size and Depth of Vocabulary Knowledge: What the research shows. *Language Learning*, 64, 913–951.
- 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. doi:10.1075/scl.45.12sch
- Sinclair, J. (2004). *Trust the Text: Lexis, corpus, discourse*. London: Routledge.
- Sun, Y.-C., & Wang, L.-Y. (2003). Concordances in the EFL Classroom: Cognitive approaches and collocation difficulty. *Computer Assisted Language Learning*, 16, 83–94.
- Vargo, M., & Blass, L. (2014). *Pathways 3 Reading, Writing and Critical Thinking*. Boston, MA: National Geographic Learning.
- Walker, I., & Hulme, C. (1999). Concrete Words are Easier to Recall than Abstract Words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1256–1271.
- Webb, S. (2007). The Effects of Repetition on Vocabulary Knowledge. *Applied Linguistics*, 28, 46–65.

- Webb, S., & Kagimoto, E. (2009). The Effects of Vocabulary Learning on Collocation and Meaning. *TESOL Quarterly*, 43, 55–77.
- Webb, S., Newton, J., & Chang, A. (2013). Incidental Learning of Collocation. *Language Learning*, 63, 91–120.
- West, M. (1953). *A General Service List of English Words*. London: Longman.
- Wolter, B., & Gyllstad, H. (2011). Collocational Links in the L2 Mental Lexicon and the Influence of L1 Intralexical Knowledge. *Applied Linguistics*, 32, 430–449.
- Wolter, B., & Gyllstad, H. (2013). Frequency of Input and L2 Collocational Processing. *Studies in Second Language Acquisition*, 35, 451–482.
- Wood, D. (Ed.) (2010). *Perspectives on Formulaic Language: Acquisition and communication*. New York, NY: Continuum.
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.
- Yamashita, J., & Jiang, N. (2010). L1 Influence on the Acquisition of L2 Collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly*, 44, 647–668.