

2019-01-01

## The Effect of Learning Environments on Thai Speakers' English L2 Vocabulary Depth

Sarut Supasiraprapa

Follow this and additional works at: <https://digital.car.chula.ac.th/pasaa>



Part of the [Reading and Language Commons](#)

---

### Recommended Citation

Supasiraprapa, Sarut (2019) "The Effect of Learning Environments on Thai Speakers' English L2 Vocabulary Depth," *PASAA*: Vol. 57, Article 5.

DOI: 10.58837/CHULA.PASAA.57.1.5

Available at: <https://digital.car.chula.ac.th/pasaa/vol57/iss1/5>

This Original Article is brought to you for free and open access by the Chulalongkorn Journal Online (CUJO) at Chula Digital Collections. It has been accepted for inclusion in PASAA by an authorized editor of Chula Digital Collections. For more information, please contact [ChulaDC@car.chula.ac.th](mailto:ChulaDC@car.chula.ac.th).

---

**The Effect of Learning Environments on Thai Speakers’  
English L2 Vocabulary Depth**

---

**Sarut Supasiraprapa**

Graduate School of Language and Communication  
National Institute of Development Administration (NIDA)  
118 Serithai Road, Bangkok, Bangkok, 10240 Thailand  
Email: sarut.sup@nida.ac.th

**Abstract**

Second language (L2) vocabulary depth, or how well learners know an L2 word (e.g., Meara, 1996), is a dimension of vocabulary knowledge that assists in L2 comprehension and production (e.g. Li & Kirby, 2014; Qian, 2002). The current study investigated the effects of two types of English L2 learning environments—formal English classrooms in Thailand and English exposure in an English-speaking country—on adult Thai speakers’ ( $N=29$ ) English vocabulary depth. Participants completed the Word Associates Test (Read, 1998), which measured the depth of their receptive vocabulary knowledge, and an elicitation task which required them to supply English collocations in a given context. Regression results based on both measures suggested that participants’ previous length of English education in Thailand did not significantly predict their English vocabulary depth, but their length of stay in the US, which followed their English education in Thailand, significantly predicted the depth. The findings were in line with the theoretical proposals (e.g., Ellis, 2013) and previous empirical results (e.g., Parkinson, 2015) suggesting the superiority of an L2 immersion environment over an environment where an L2

is a foreign language in promoting the depth dimension of L2 word knowledge. Based on the results, pedagogical implications are discussed.

**Keywords:** Vocabulary depth, L2 learners, Word Associates Test, collocations

### **Vocabulary Depth**

In L2 research, researchers have distinguished between two dimensions of vocabulary knowledge: breadth and depth (e.g., Meara, 1996, 2009; Nation, 2001; Schmitt & Meara, 1997). The former constitutes the basic dimension of lexical competence and refers to the size or the number of words of which a learner has at least some knowledge about the meaning. By contrast, depth pertains to how well learners know a word and, according to Read (2004), has been conceptualized in various ways. As Schmitt (2014) pointed out, however, the most common conceptualization is vocabulary depth as the degree to which L2 words are linked to related words in the mental lexicon, or L2 speakers' ability to link an L2 word to and distinguish it from related words. Thus, vocabulary depth is typically measured from L2 speakers' associative behaviors, especially the degree of nativelikeness in word association tasks.

According to Schmitt (2014), this conceptualization is the most widely adopted due to the availability of Read's (1993, 1998) Word Associates Test (WAT), which measures English lexical network knowledge of non-native English speakers (NNSs). In this test, NNSs are instructed to associate each given adjective (e.g., *bright*) with synonymous or semantically related adjectives (e.g., *clever*, *shining*) and with noun collocates (e.g., *color*). Thus, to supply correct answers, test takers need to know that an English adjective may have a range of meanings/ synonyms and know what nouns it can collocate with. Despite a concern about the WAT's validity (i.e., it may invite guessing), which makes the test unsuitable as a high-stake test, previous research has demonstrated that the WAT was generally reliable and useful for

research and classroom applications (Schmitt, 2010; Schmitt, Ng, & Garras, 2011; Schoonen & Verhallen, 2008). Moreover, there is evidence that the breadth and the depth, the latter being measured by the WAT or a modified version of the WAT, each significantly accounted for NNSs' English reading and listening comprehension (Li & Kirby, 2014; Qian, 1999, 2002; Stæhr, 2009). Empirical evidence has also demonstrated that WAT scores significantly predicted NNSs' success in inferring English word meanings during reading (Nassaji, 2006) and the quality of their English summary writing (Li & Kirby, 2014).

### **Collocation Use and Native-Like Selection as Part of L2 Vocabulary Depth**

Besides research adopting the WAT, given the conceptualization of L2 vocabulary depth as nativelike L2 word associations, also of relevance is L2 research on NNSs' knowledge or use of English collocations (see e.g., Henriksen, 2013). In comparison to the operationalization of vocabulary depth as WAT scores, therefore, collocation knowledge constitutes one part of vocabulary depth because the WAT measures knowledge of both collocates and synonyms of a word.

It should be pointed out that L2 research researchers have defined the term *collocation* based on two main approaches. First, in the phraseology approach (e.g., Cowie, 1998; Nesselhauf, 2005), a collocation is a word sequence (e.g., *strong tea*) in which at least one word conveys a figurative meaning (e.g., *strong*) and at least one word expresses a literal meaning (e.g., *tea*). A collocation has a transparent meaning, but component word substitution is subject to arbitrary restriction (e.g., *powerful tea* is unidiomatic although *powerful* and *strong* are semantically related). In the second approach, or the statistics approach (e.g. Gries, 2010; Sinclair, 1991), collocations are word sequences with an above-chance possibility of co-occurrence in a corpus. Collocations may be sequences in which every word is used literally (e.g., *green tea*), sequences in which at least a word is used figuratively (e.g., *strong tea*), or idioms (e.g., *kick the bucket*), which have a meaning not

derivable from word parts. These therefore set apart the two approaches to defining collocations. Similar to many recent L2 collocation studies (e.g., Wolter & Gyllstad, 2013; Wolter & Yamashita, 2018), the current study adopted the statistics approach because it allows for objective identification of collocations, which is less likely to be affected by differences in opinions of native English speakers (NSs).

Previous empirical studies investigating NNSs' collocation knowledge typically compared collocations produced by NNSs in written essays against those produced by NSs. Such studies have reported some consistent findings. For example, adult NNSs have limited knowledge of English collocations, may use malformed phrases containing words familiar to them (e.g., *do a mistake*), and may underuse or overuse some collocations (e.g., Granger & Bestgen, 2014; Nesselhauf, 2005). Some other studies used elicited production tasks, such as translation tasks, and reported similar results. For example, while semantically related adjectives are not always interchangeable as noun collocates, Farghal and Obiedat (1995) reported that, in writing, ESL learners may use an inappropriate semantically related adjective (e.g., *heavy tea*) instead of a correct adjective (e.g., *strong tea*).

Despite such useful findings based on NNSs' collocation production, as Foster, Bolibaug, and Kotula (2014) argued, such results may not reflect collocation knowledge that NNSs have but cannot use productively. Therefore, according to these researchers, another way to measure L2 word association is to ask NNSs to judge whether given English word combinations are acceptable. Foster et al. (2014) asked NNSs to identify unidiomatic English phrases embedded in short paragraphs (e.g., *The boys tried many efforts to reach the ball, but did not get success because the hole was so deep*) and referred to correct identification as native-like selection. The results revealed that only NNSs who started living in an English-speaking country before the age of 12 were able to perform similarly to NSs. A judgement task was also used in a study by Sonbul (2015), in which adults NNSs at a UK university gave typicality rating scores to adjective-noun

collocations (e.g., *huge crowd*) and unidiomatic phrases containing synonymous adjectives (e.g., *grand crowd*). Based on the results, these NNSs rated more frequent collocations as being more typical. However, in comparison to NSs, the NNSs gave relatively high rating scores to non-collocations. In another study, Siyanova and Schmitt (2008) used an acceptability judgement task and reported similar findings based on another group of adult NNSs at a UK university.

To conclude, the most common operating definition of vocabulary depth in empirical L2 research is the degree to which L2 words are linked to related words in the mental lexicon, or the learners' ability to link a word to and distinguish it from related words. The WAT score is the most widely used measure of NNSs' English vocabulary depth, but related measures also include NNSs' English collocation use and their ability to judge whether given English phrases are idiomatic.

### **Vocabulary Depth and L2 Input Exposure**

Based on these measures, to acquire vocabulary depth, NNSs need to establish (1) a link between a word and its multiple meanings/ synonyms and/or (2) a link between a word and its collocates. In light of the L2 literature, acquiring such depth poses several challenges for NNSs and entails a great deal of English exposure. First, associating a word with multiple meanings to its synonyms is not easy because the meanings of such a word are contextually dependent. For example, while *bright* can mean *clever* (e.g., *bright children*), it can also mean *shining* (e.g., *bright stars*) or *full of hope and success* (e.g., *bright future*). Because learning a word requires repeated exposure to the word (Nation, 2001; Schmitt, 2008; Webb, 2008), to know different meanings/ synonyms of a word, NNSs need to repeatedly encounter examples of each different meaning in contexts. The amount and diversity (i.e., variety of contexts) of English input is therefore pivotal. However, this does not mean that explicit L2 vocabulary instruction offers no advantage. As Schmitt (2008) pointed out, such instruction can help NNSs establish initial L2 word form-

meaning mappings, but contextualized aspects of vocabulary knowledge have to be acquired through substantial L2 input exposure. Given these observations, therefore, one possible reason for a weak or no link between an L2 word and its synonyms is insufficient exposure to the word in different contexts.

Regarding a link between an L2 word and its collocates, researchers informed by usage-based approaches to L2 acquisition—notably Ellis (2011, 2013)—account for how such a link is established as part of L2 speakers' knowledge. According to Ellis (2011, 2013), such acquisition can be ascribed to the human cognitive ability of chunking, which allows a word string (e.g., *bright future*) to be registered in learner memory and establishes a sequential relation between constituent words (e.g., *bright* and *future*). This relation will become stronger if L2 speakers are later exposed to the collocation repeatedly in L2 input. Moreover, L2 collocation acquisition may be facilitated by L2 instruction or practice that draws learners' attention to collocations; that is, such interventions help register a sequence in a learner's memory before the relation between component words in a collocation is strengthened through subsequent collocation exposure. Similarly, Schmitt (2008) contended that collocation knowledge constitutes a contextualized aspect of vocabulary knowledge, which is difficult to teach explicitly and is more likely to be acquired through massive L2 input exposure. Although not every L2 researcher shares Ellis' view (Wray, 2002), numerous empirical studies to date have lent support to his claim (e.g., Durrant & Schmitt, 2010; Wolter & Yamashita, 2018), suggesting that L2 collocation is possible and that frequency of collocation exposure is vital for L2 collocation acquisition.

In sum, the L2 literature suggests that a great deal of English input is necessary for an increase in NNSs' L2 vocabulary depth as measured by nativelike word associations. Therefore, the types of English learning environment where NNSs live play a significant role in their L2 vocabulary depth acquisition.

## **L2 Learning Environments and Vocabulary Depth**

According to Muñoz (2008), L2 learning can be classified into two main types: *naturalistic second language learning* and *foreign language learning*. The former takes place in an L2 immersion environment, in which L2 learners have an opportunity to have *significant exposure* to L2 input—that is, “they are able to carry out a variety of speech acts over a wide range of situations and topics, and to participate in social settings effectively dominated by the L2.” (p.585). In contrast, a foreign language learning environment has some or all of the following characteristics. First, L2 instruction is limited to around 2-4 sessions per week, each lasting about 50 minutes. Moreover, the L2 input that learners receive is mainly from one speaker, which is typically an L2 teacher. However, not all L2 teachers communicate in the L2 in the classroom, and teachers’ general and oral L2 proficiency varies greatly. Further, the L2 is not the language of communication among students and is not spoken outside the English classroom. Learners therefore usually receive none or not substantial nativelike L2 input. Based on a review of empirical studies, Muñoz (2008) further noted that, in the case of NNSs who learn English in non-immersion classrooms in a foreign language learning environment, there is no evidence that an earlier starting age of English learning leads to a superior long-term learning outcome.

In the current study, which focuses on English as an L2, the naturalistic and the foreign language learning environments will be referred to as English-as-a-second-language (ESL) and the English-as-a-foreign-language (EFL) contexts, respectively. It has been observed that English classrooms alone are unlikely to provide sufficiently rich L2 input for learners to develop nativelike networks of L2 word associations (Hoey, 2005). Moreover, regarding collocations, NNSs are unlikely to have multiple encounters with English collocations in non-immersion contexts, even when the component words are frequent, leading to unsuccessful English collocation acquisition (Boers, Demecheleer, Coxhead, & Webb, 2014). Given these observations, an ESL



environment is thus more likely to promote NNSs' English vocabulary depth than an EFL environment. Therefore, their English vocabulary depth may increase as a function of the amount of time they spend in an ESL context. On the other hand, such depth may not correlate with the amount of time they study English in non-immersion classrooms in an EFL context. These possibilities are of central interest in the current study.

To date, empirical evidence has indeed suggested that L2 vocabulary depth increases as a result of living in an ESL environment. These studies were based on investigation of NNSs' collocation use in writing or speaking. For example, Li and Schmitt (2009) found that, after one year, a participant who was a graduate student in the UK not only acquired many new common English phrases (e.g., *on the other hand*) but also used such phrases more appropriately. In another study, Crossley and Salsbury (2011) reported that, the longer six adult NNSs stayed in the US, the more likely that they said phrases frequently used by NSs in casual conversations. More recently, focusing on the English motion construction (e.g., *go to Mexico, come down*)—in which the verb *go* and *come* are most common—Li, Eskildsen, and Cadierno (2014) observed one adult NNS over 3.5 years and found that, as the NNS spent more time in the US, he used this construction more productively and associated these verbs with new collocates, such as new propositions (e.g., *go out, come into*).

There is also empirical evidence suggesting the superiority of an ESL environment over an EFL environment in promoting L2 word associations. For example, Nesselhauf's (2005) analysis of essays written by NNSs who had studied English in an EFL environment for 5-17 years revealed that "more years of [classroom English] teaching apparently fail to lead to an increased use of [English] collocations." (p.235). In another study, Groom (2009) found that, compared to an EFL environment, an ESL environment better promoted correctness and greater variety of NNSs' collocation use in writing. Moreover, a written corpus analysis by Parkinson (2015) revealed that NNSs in an ESL environment used noun-noun collocations (e.g., *electricity shortage*) significantly

more accurately than NNSs with a matched English proficiency level in EFL settings. Finally, Foster et al. (2014) reported that, based on the ability to correctly identify unidiomatic English word combinations embedded in a context story, NNSs in an ESL environment significantly outperformed similarly proficient NNSs who lived in an EFL context.

### **The Current Study**

Focusing on the effect of learning environments on English L2 vocabulary depth, the current study was motivated by a few gaps in empirical research. First, while the WAT score is the most common measure of such depth, to the best of the researcher's knowledge, no study has focused on the effects of different English learning environments on these scores. Moreover, while previous research has investigated the possible effects of English exposure on NNSs' collocation use in writing or speaking (e.g., Crossley & Salsbury, 2011; Parkinson, 2015), as Foster et al. (2014) pointed out, production-based measures may not shed light on knowledge of L2 word associations that NNSs have but cannot draw on productively. Therefore, using WAT scores, the current study addressed this limitation because the WAT is a receptive knowledge test. Moreover, as discussed, the WAT provides a more comprehensive assessment of English vocabulary depth because the test assesses NNSs' knowledge of both collocations and synonyms. In addition, the current study used a collocation elicitation task to assess NNSs' vocabulary depth since several previous studies investigating the effect of an English learning environment on L2 collocation use over time were based only on a few participants or on descriptive statistics such as the mean number of collocations produced (e.g., Crossley & Salsbury, 2011; Groom, 2009; Li & Schmitt, 2009), casting doubt on the generalizability of the findings. The research questions that guided the current study were as follows:

1. Can length of English education in formal classrooms in an EFL environment predict adult NNSs' vocabulary depth?

2. Can length of stay in an English-speaking country predict adult NNSs' vocabulary depth?

If the results suggest that, unlike an ESL environment, the Thai EFL environment does not help promote the acquisition of English vocabulary depth, in light of Muñoz's (2008) observations, the results will therefore suggest a necessity for an increase in the amount, quality, and variety of English input for Thai EFL learners.

## **Method**

### ***Participants***

Participants were 29 adult Thai learners of English living in Bangkok, Thailand (Male = 17, Female = 12, age range 21-39, mean age = 30.48,  $SD = 4.74$ ). They were from a variety of educational and professional backgrounds (e.g., undergraduate student, marketing officer, business owner, university professor, researcher). They all began learning English as an L2 in formal classrooms in Thailand, an EFL context. The participants reported that the number of years they had studied English in such classrooms ranged from 7 to 17 years ( $M = 12.17$ ,  $SD = 2.83$ ). No participants studied in an English immersion environment at a young age—although some of them started learning English early, their English teachers were Thai and mainly spoke Thai in English classrooms. Thus, none can be classified as early learners of English, who were immersed in an English-speaking environment (i.e., an international school in Thailand) at a young age (e.g., before the age of 12, the age criterion used by Foster et al., 2014). The number of years the participants had previously spent in the US ranged from 0 to 13 ( $M = 3.2$ ,  $SD = 3.99$ ). For participants with experience of living in the US, they went there after their EFL education in Thailand to pursue an undergraduate degree and/or a graduate degree. Prior to their stay in the US, these participants lived in Thailand and had never lived in an English-speaking country. Only learners who had returned to Thailand for no more than three years were included so that participants were

similar in terms of how long they had returned to their home country.<sup>1</sup>

Participants in the current study had a minimum paper-based TOEFL score of 550, or rough equivalents in other test formats. Based on previous studies, these participants can be characterized as being proficient enough to study in an English-speaking university environment (e.g., Durrant & Schmitt, 2009; Sonbul, 2015). Note that two participants, who were undergraduate students, had never lived in the US, but they took the TOEFL as part of their preparation to study abroad.

## Materials

### *Word Associates Test*

Read's (1998) WAT, available on [www.lex tutor.com](http://www.lex tutor.com), was used to assess participants' receptive vocabulary depth. An example of WAT test items is shown below.

bright

clever	famous	happy	shining	color	hand	poem	taste
--------	--------	-------	---------	-------	------	------	-------

In this item, the given adjective is *bright*. Participants were informed that the four answer choices in the left box are words that may help explain the meaning of the given adjective, while the four words in the right box are nouns that may collocate with the adjective. According to Read (1998), the reason why both synonyms and the noun collocates of each target adjective were included as answer choices was because an adjective can have both a paradigmatic relationship and a syntagmatic relationship with other words, respectively. Participants had to select four answer choices and were informed that, while all test items had four correct answers, the number of correct answers on the left and the right boxes varied throughout the test. In total, the test consisted of 40 items, with a perfect score of 160. The test was copied from the website and presented to participants in a paper

format to obviate a need for a computer access and an internet connection during data collection.

### ***Collocation Elicitation Task***

Adjective-noun collocations were chosen as target collocations because they can be problematic even for advanced ESL learners (e.g., Sonbul, 2015). The selection of target collocations followed a few major steps. First, as in Boers et al.'s (2014) study, candidate collocations were sampled from the textbook *English Collocation in Use* (McCarthy & O'Dell, 2005). Moreover, because participants in the current study were relatively proficient in English, additional candidate collocations were sampled from *Collocation in Use – Advanced* (McCarthy & O'Dell, 2008), a textbook in the same series for more advanced learners. Based on an analysis of several corpora (i.e., Cambridge International Corpus of Written and Spoken English, the Cambridge and Nottingham Corpus of Discourse in English, and the Cambridge Learner Corpus), the textbook authors maintained that collocations in these books have high communicative usefulness and are potentially problematic for NNSs. Another consideration for the collocation selection was that constituent words in target collocations should be familiar to the participants; otherwise, if they demonstrated a lack of collocation knowledge, it would be unclear whether they did not know a collocation or did not know the component words. In total, 30 collocations were initially selected, but due to a difficulty in incorporating all of them into the collocation elicitation task described in the next section, 28 were chosen as target collocations.

Second, as noted, the current study adopted a statistics approach to defining collocations (e.g., Gries, 2010; Sinclair, 1991), the 28 target collocations were therefore entered into the Corpus of Contemporary American English (COCA; Davies, 2013) so that their mutual information (MI) scores and frequency were obtained. The frequency of each collocation was then used to calculate a t-score using an online calculator developed based on the guidelines provided by Stubbs (1995) at

<http://www2.lael.pucsp.br/corpora/association/calc.htm>.

T-scores and MI scores are corpus-based measures indicating whether the possibility of co-occurring words is above chance, and these measures tend to highlight collocations consisting of low and high frequency words, respectively (e.g., Evert, 2007).<sup>2</sup> Because a word sequence with an MI score of three or above and/or a t-score of two or above is a collocation (Durrant & Schmitt, 2009; Stubbs 1995) and the 28 target sequences met this criterion, they can be considered collocations from a statistics perspective. The collocations, their MI scores, and t-scores, are shown in the Appendix in the order of their appearance in the task.

Subsequently, a native American English speaker with experience in teaching English to NNSs was hired to construct a short story in which the target collocations were embedded. To create context sentences, the story writer was instructed to use words that should be known to the participants. The story was also based on a topic that should be familiar to them (i.e., vacation). Below are examples of the context sentences, with embedded target collocations in bold.

It was the middle of April, and the end of the school year was quickly approaching. Jessica and Keith loved summer for one (1) **simple reason**: vacation. Every summer after school gets out, Jessica and Keith take a special vacation to a new place. Jessica has a (2) **strong preference** for city vacations, but Keith prefers outdoor adventure trips. This is a (3) **big problem** when they try to decide where to travel, and because of their (4) **strong opinions**, they usually end up having a (5) **heated debate** about their vacation destination.

Once the story was constructed, the adjective in each target collocation (e.g., *simple*) was replaced with an adjective that is synonymous or semantically close but produced an unidiomatic sequence (e.g., *easy*). The reason was because previous studies have suggested that even relatively advanced English learners may

have difficulty distinguishing semantically related adjectives that are not always interchangeable as a noun collocate (e.g., Farghal & Obiedat, 1995). T-scores and MI scores for the unidiomatic sequences (e.g., *easy reason*) were then calculated, and the calculation revealed that the unidiomatic sequences failed the MI and t-score criterion. This confirmed that the unidiomatic sequences were indeed not collocations. Examples of final sentences in this task are shown below:

It was the middle of April, and the end of the school year was quickly approaching. Jessica and Keith loved summer for one (1) **easy reason**: vacation. Every summer after school gets out, Jessica and Keith take a special vacation to a new place. Jessica has a (2) **heavy preference** for city vacations, but Keith prefers outdoor adventure trips. This is a (3) **large problem** when they try to decide where to travel, and because of their (4) **powerful opinions**, they usually end up having a (5) **burning debate** about their vacation destination.

In this task, participants were instructed to read the story and consider whether the 28 unidiomatic adjective-noun sequences, which were in bold, sounded natural. If any sounded unnatural, participants should cross out the adjective and provide a correction while trying to keep the noun and the meaning the same. To get the perfect score of 28, the participants had to correct all the 28 bolded phrases. Therefore, this task combined two characteristics of word association tasks in previous research: judgement of whether a given word combination was acceptable (Foster et al., 2014) and NNSs' ability to supply a correct collocation in a given context (e.g., Farghal & Obiedat, 1995). The instruction did not specify the number of incorrect sequences or whether all the bolded sequences sounded unnatural. The instruction was also translated into Thai to ensure the participants' understanding.

### ***Background Questionnaire***

A background questionnaire was used to collect participants' information such as their age, latest TOEFL score, educational and professional background, the age at which they started learning English, the length of their EFL education in Thailand, and the amount of time they had previously spent in an English-speaking country.

### **Procedures**

Each participant completed the WAT and the elicitation task in a quiet room at a location which is convenient for the participant (e.g., their offices). Participants completed the elicitation task before the WAT because the WAT contained many more test items, and based on a pilot test, participants may become fatigued after WAT test completion. Participants were given 20 and 25 minutes to do the two respective tests because the pilot test also indicated that these should be sufficiently long for them. After the WAT, they completed the background questionnaire. In total, these procedures lasted approximately 50 minutes.

### **Scoring**

Participants' answers to the paper-based WAT were entered for online scoring at [www.lex tutor.com](http://www.lex tutor.com). Regarding the elicitation task, another native American English speaker, who was an English instructor at a large university in Bangkok, was hired to score the answers based on an answer key containing the target answers. To get a score, the participants must be able to identify that a bolded phrase in the story was unidiomatic and replace the wrong adjective with a correct one. However, there were some cases in which the participants replaced a wrong adjective (e.g., *large problem*) with a possible adjective (e.g., *huge problem*) that was not a target (e.g., *big problem*). In such cases, the researcher consulted with both the NS who created the story and the NS who taught at a university Bangkok for mutual agreement. The researcher also checked with COCA if the non-target yet possible answer passed the MI score and t-score criterion (i.e., the answer was a collocation based on the statistics approach). A non-target yet possible answer received a score if both NSs agreed that the



answer was acceptable and the adjective-noun sequence passed the corpus-based criterion.

Results

Table 1 shows scores from the WAT and the elicitation tasks. The participants’ mean WAT score was 131.93, or approximately 82% of the full score, while the mean score from the elicitation task was 15.28, or about 55% of the full score. Therefore, it seemed that the elicitation task was more challenging than the WAT to the participants.

Table 1. Descriptive statistics of WAT and collocation elicitation task scores (N=29)

	M	SD	Min	Max	Full score
WAT scores	131.93	10.99	108	149	160
Collocation elicitation task scores	15.28	4.41	7	26	28

A detailed investigation of the answers from the elicitation task further revealed that the NNSs may consider some unidiomatic sequences (e.g., *easy reason*, *large importance*) acceptable—that is, they did not cross out some bolded phrases in the story. Moreover, as Table 2 shows, on some occasions, participants correctly crossed out unidiomatic phrases (e.g., *big details*) but supplied other unidiomatic adjective-noun combinations as answers (e.g., *large details* or *deep details*). These unidiomatic phrases contained incorrect adjectives which are synonymous or semantically related to adjectives which should have been used in the given context (e.g., *great details*).

Table 2. Examples of incorrect answers in the elicitation task

Bolded phrases	Target elicited answers	Examples of incorrect answers
big pleasure	great pleasure	high pleasure
big detail	great detail	large/deep detail
heavy preference	strong preference	big/large/great preference
strong rains	heavy rains	hard rains
durable tradition	longstanding tradition	continuous tradition

As pointed out earlier, for these participants, their EFL education preceded their ESL exposure. Figures 1a and 1b show a scatterplot between WAT scores and their EFL education length and a scatterplot between WAT scores and their ESL exposure length, respectively. A Pearson's correlation analysis indicated a non-significant negative correlation between WAT scores and length of EFL education,  $r = -.25$ ,  $p = .197$ , with a small-to-medium effect size,  $R^2 = .06$ , based on Cohen's (1992) guidelines. On the other hand, the number of years participants spent in the US correlated positively and significantly with WAT scores,  $r = .54$ ,  $p = .002$ , with a large effect size,  $R^2 = .29$ . This suggested that the longer the participants lived in the US, the higher WAT scores they obtained.

Figure 1. Scatterplots between WAT scores and (a) EFL education length and (b) ESL exposure length in years

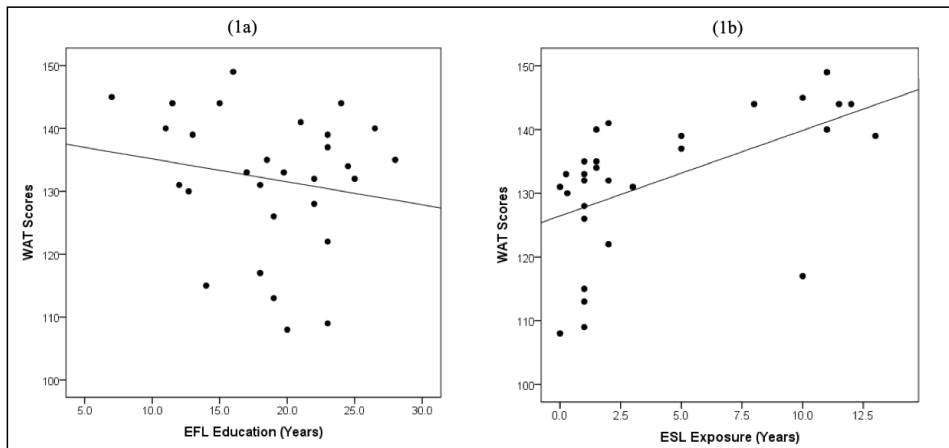
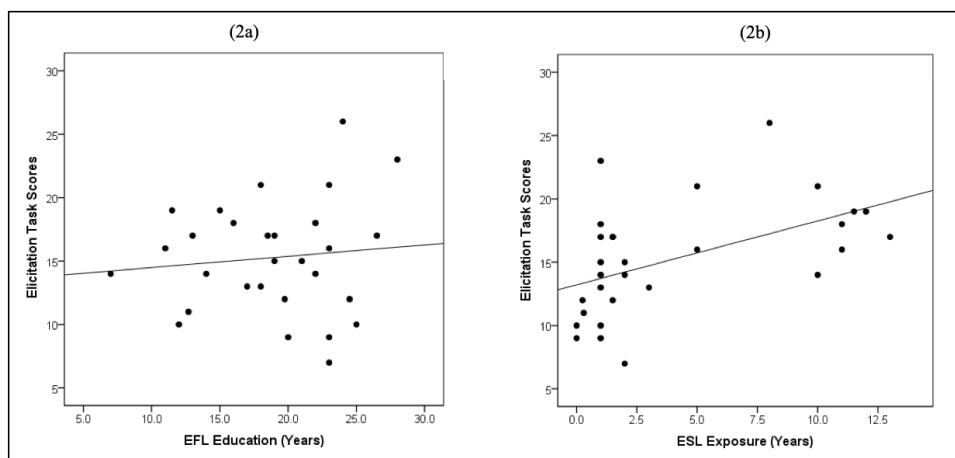


Figure 2. Scatterplots between collocation elicitation task scores and (a) EFL education length and (b) ESL exposure length in years



Figures 2a and 2b show a scatterplot between collocation elicitation task scores and EFL education length and a scatterplot between the scores and ESL exposure length, respectively. The correlation between elicitation scores and EFL education length was positive but non-significant, with a negligible effect size,  $r = .08$ ,  $p = .675$ ,  $R^2 = .01$ . In contrast, there was a significant positive relationship between elicitation task scores and ESL exposure length with a large effect size,  $r = .50$ ,  $p = .005$ ,  $R^2 = .25$ , indicating that the longer the participants lived in the US, the higher scores they obtained.

To answer the two research questions, linear regression analyses were performed. Table 3 shows the results when the dependent variable was WAT scores, while the length of EFL education and ESL exposure were predictors. Moreover, because the participants had received EFL education before going to the US, an interaction between EFL education length and ESL exposure length was additionally included to shed light on whether the effect of ESL exposure also depended on the length of their previous EFL education (e.g., the effect of ESL exposure on WAT scores may be stronger if participants had previously received longer EFL education). However, as Cohen, Cohen, West,

and Aiken (2003) suggested, including an interaction term between two continuous variables may lead to a problem of collinearity in a regression model; therefore, to avoid this problem and in line with Cohen et al.'s (2003) suggestion, each predictor was centered around its mean in the analysis.<sup>3</sup>

Table 3. Linear model of predictors of WAT scores

	<i>B</i>	<i>SE B</i>	95% CI	$\beta$	<i>p</i>
Constant	131.69	1.97	(127.63, 135.76)		<.001***
Length of EFL education (years)	-0.16	0.71	(-1.62, 1.30)	-0.04	.822
Length of ESL exposure (years)	1.24	0.48	(0.25, 2.23)	0.50	.016*
EFL*ESL	-0.05	0.17	(-0.41, 0.30)	-0.06	.765

*Note.*  $R^2 = .30$

As Table 3 shows, EFL education length was not a significant predictor of WAT scores ( $p = .822$ ), while the length of ESL exposure significantly predicted the scores ( $p = .016$ ). The positive coefficient suggested that the longer the participants had lived in the US, the higher WAT scores they obtained. The interaction between EFL education length and ESL exposure length was not significant ( $p = .765$ ). Variance inflation (VIF) scores for all predictors were below 10, suggesting that collinearity was not a problem in this model (Loewen & Plonsky, 2016).

Table 4 shows the regression results when elicitation task scores were the outcome measure. The effect of EFL education length did not reach significance ( $p = .102$ ), but the length of ESL exposure significantly predicted the scores ( $p = .001$ ). The positive regression coefficient suggested that elicitation task scores increased as a function of ESL exposure length. Moreover, the interaction between EFL education length and ESL exposure length was not significant ( $p = .394$ ). Based on VIF scores, collinearity was also not a problem in this model.

Table 4. Linear model of predictors of collocation elicitation task scores

	<i>B</i>	<i>SE B</i>	95% CI	$\beta$	<i>p</i>
Constant	15.54	0.76	(13.97, 17.11)		<.001***
Length of EFL education (years)	0.47	0.27	(-0.10, 1.03)	0.30	.102
Length of ESL exposure (years)	0.68	0.19	(0.30, 1.07)	0.68	.001**
EFL*ESL	0.06	0.07	(-0.08, 0.20)	0.16	.394

Note.  $R^2 = .36$

In sum, the results suggested that, for these adult Thai speakers, the length of EFL education did not significantly predict their English vocabulary depth. On the other hand, the number of years they had spent in the US significantly predicted such depth; a longer ESL exposure brought about a vocabulary depth increase.

## Discussion

Drawing on the commonly adopted definition of vocabulary depth as the extent to which L2 words are linked to other words in the mental lexicon (Meara, 1996, 2009), the current study investigated the effect of EFL education length and ESL exposure length on adult Thai speakers' English vocabulary depth. Regarding the first research question, the results suggested that the length of previous EFL education in formal classrooms in Thailand, a non-English immersion context, did not predict the participants' vocabulary depth, whether the depth was measured with scores from the WAT or the adjective-noun collocation elicitation task. These participants cannot be classified as early learners of English, who learned English in an immersion environment at a very young age (e.g., before the age of 12, the criterion used by Foster et al., 2014). The EFL environment where they lived fitted the characteristics of a foreign language learning environment described by Muñoz (2008). That is, their exposure to English was limited in terms of source, quantity, and quality due to various reasons—including the limited amount of English instruction (typically two to four 50-minute sessions per week), a lack of the variety of L2 input sources, English teachers' use of

Thai in classrooms, a variability in English teachers' general and oral English proficiency, and the fact that English was not the language of communication among the participants outside English classrooms.

Such results are in line with the observation that, in an EFL setting, there is no clear evidence that an earlier starting age (i.e., longer length) of L2 learning in a non-immersion environment guarantees a superior long-term L2 learning outcome (Muñoz, 2008). In addition, because vocabulary depth results from massive L2 exposure input—including repeated exposure to a word, which may have multiple meanings, in different contexts (e.g., Nation, 2001; Schmitt, 2008) and multiple collocation encounters (e.g., Ellis, 2011, 2013; Schmitt, 2008)—the results are compatible with the observations that English input in an EFL setting is unlikely to be sufficient in terms of amount and diversity for NNSs to develop nativelike networks of English word associations (Boers et al., 2014; Hoey, 2005; Nesselhauf, 2005).

With regard to the second research question, the results indicate that NNSs' vocabulary depth, measured as WAT scores or collocation elicitation task scores, increased as a consequence of a longer length of stay in an English-speaking country. After their EFL education in Thailand, most participants in the current study went to the US to pursue an undergraduate and/ or a graduate degree and subsequently returned to their home country. For example, one participant, who obtained a relatively high WAT score of 149, went to the US after his high school education in Thailand and stayed there for more than 10 years to pursue an undergraduate and two graduate degrees. Such an ESL context, as Muñoz's (2008) observed, provided learners with significant exposure to English input and a variety of contexts of English use and interaction. The results from the current study therefore corroborate previous empirical findings that L2 input in an immersion context brought about an increase in English L2 vocabulary depth, as measured by collocation use in writing or speaking (e.g. Crossley & Salsbury, 2011; Li, Eskildsen, & Cadierno, 2014; Li & Schmitt, 2009). The findings also support

previous empirical results suggesting the superiority of an ESL environment over an EFL environment in promoting vocabulary depth, as indicated by productive English collocation use (Parkinson, 2015) or correct identification of malformed English phrases (Foster et al., 2014).

As pointed out, the current study included both the WAT and the collocation elicitation task due to the differences in the operationalization of vocabulary depth in previous empirical studies. The WAT is the most common measure of such depth and is a receptive vocabulary knowledge test; thus, it can potentially provide insight into vocabulary knowledge which learners have but may not be able to draw on productively. Also, to the best of the researcher's knowledge, no previous studies had investigated the relationship between English learning environments and WAT scores, and therefore the current study filled in this research gap. The collocation elicitation task was additionally included because several previous studies using production-based measures of vocabulary depth suggested that such depth increases as a result of longer ESL exposure, but these studies relied only on a small number of participants and/ or descriptive statistics (Crossley & Salsbury, 2011; Groom, 2009; Li & Schmitt, 2009). Therefore, the findings from the current study seemed to indicate the generalizability of the results from those previous studies.

In terms of pedagogical implications, based on the results, to help Thai learners of English develop English vocabulary depth, curriculum developers or English teachers in Thailand therefore need to make greater efforts to provide Thai EFL learners with a greater amount and variety of English input, which includes repeated exposure to a word in different contexts and multiple collocation encounters (e.g., Ellis, 2011, 2013; Schmitt, 2008). Thus, incorporating additional authentic English input (e.g., listening materials) into English classrooms or requiring students to read or listen to English in supplementary learning materials outside classes may be helpful. Certainly, such materials need to be carefully designed so that the materials are appropriate for students' English proficiency levels and specific target learning

outcomes (e.g., target words or collocations) can be expected. Moreover, in light of Muñoz's (2008) discussion, increasing the English proficiency of English teachers in Thailand should be beneficial because these teachers are the main sources of English input for students. This may pose a practical challenge; empirical research has revealed that even some Thai English teachers acknowledged that they were not proficient in English and that they lacked confidence when communicating in English (e.g., Prapaisit de Segovia & Hardison, 2008). However, this is a challenge that needs to be overcome.

While underscoring the significance of L2 input exposure, the current study does not argue that explicit instruction of L2 words and collocations is of no benefit. As discussed, explicit teaching of synonyms or collocations can be useful but must be followed by repeated exposure to target words and collocations (Ellis, 2011, 2013; Schmitt, 2008). It has also been observed that when NNSs encounter English collocations, they may not pay attention to such collocations, especially when collocations contain familiar words (e.g., Boers et al., 2014; Laufer & Waldman, 2011). In fact, this possibility may explain why even participants in the current study who had lived in the US for years did not obtain the perfect score in the collocation elicitation task, even though the constituent words in the target collocations should be familiar to them. Therefore, drawing NNSs' attention to English synonyms and collocations may help facilitate word and collocation acquisition, but such attention drawing should be followed by abundant input with salient examples of target words and collocations in contexts. Moreover, collocation production practice can be helpful. According to Ortega (2015), L2 production may promote conscious processes that help support chunking of collocated words and automatization of collocation use. Indeed, some studies have suggested possible facilitative roles of L2 collocation instruction and practices in L2 collocation acquisition (e.g., Laufer & Girsai, 2008; Webb & Kagimoto, 2009).

The results additionally suggested that negative feedback which indicates that a phrase NNSs produce is unidiomatic may



be necessary for elimination of their false hypotheses regarding L2 word combinations. In the collocation elicitation task, the participants may consider some malformed adjective-noun combinations acceptable (e.g., *easy reason*, *large decision*) or correctly identified that a combination is unacceptable but supplied a wrong adjective-noun combination (Table 2). These support the previous observations that NNSs may not know that semantically-related adjectives are not always interchangeable as noun collocates (e.g., Farghal & Obiedat, 1995). Because an inappropriate use of collocations, such as adjective-noun collocations, does not necessarily cause a communication breakdown (Laufer & Waldman, 2011), communicating in English outside classrooms may not help NNSs become aware that the phrases they use are unidiomatic. Therefore, negative feedback in English L2 classrooms may be essential because such feedback helps eliminate a false L2 hypothesis (Larsen-Freeman, 2003). This in turn means that EFL classrooms should provide ample opportunities for collocation production so that relevant feedback from teachers can be provided.

### **Limitations**

The current study contained some limitations that future studies may address. First, regarding the elicitation task, the possible influence of the participants' L1 was not strictly controlled for. In L2 research (e.g., Yamashita & Jiang, 2010), L2 collocations consist of two types: congruent and incongruent. The former type has a direct word-for-word L1 translation equivalent, while the latter type does not. Because evidence has suggested that congruent L2 collocations is easier for L2 speakers to acquire (e.g., Wolter & Gyllstad, 2013; Yamashita & Jiang, 2010), future studies may investigate whether a length of L2 exposure differently affects congruent and incongruent L2 collocation acquisition. Moreover, the elicitation task in the current study may have underestimated the participants' collocation knowledge. The potential benefit of including the malformed adjective-noun sequences in the story was that the researcher maximized the possibility that the participants would supply collocations with the target meanings. Moreover, based on previous research, correct

identification of malformed phrases was indicative of vocabulary depth (Foster et al, 2014). However, because all the bolded phrases in the story were unidiomatic but the participants had to identify whether these sequences were acceptable, they might have doubted their intuition and did not perform as well as they should have. Thus, to assess NNSs' English collocation knowledge, future studies investigating the effects of L2 learning environments may use other types of elicitation tasks which did not require NNSs' judgement of malformed phrases.

## **Conclusion**

The current study investigated the effects of English learning environments on adult Thai speakers' English vocabulary depth, or the degree of native-like word associations. The results suggested that participants' previous length of stay in the US significantly predicted their vocabulary depth, but the length of their EFL education in formal English classrooms in Thailand did not. These results supported previous observations and empirical findings that, in comparison to an L2 immersion environment, a foreign language learning environment—where L2 input is limited in terms of quantity, quality, and variety—is less likely to promote an increase in L2 vocabulary depth. The results suggested a necessity for incorporation of more English input into EFL courses in Thailand, whether in class or in supplementary out-of-class materials, and a need for an improvement in the English proficiency of Thai EFL teachers. Finally, the results suggested that explicit classroom instruction and practice may help facilitate NNSs' English word associations and eliminate their false hypotheses regarding non-nativelike L2 word associations.

## **The Author**

Sarut Supasiraprapa is a lecturer in the Graduate School of Language and Communication at the National Institute of Development Administration in Bangkok. His work has appeared in *TESOL Quarterly*, *Bilingualism: Language and Cognition*, and *Applied Psycholinguistics*. His research interests include usage-based and

instructed second language acquisition, psycholinguistics, and research methods.

## Notes

<sup>1</sup> To maximize the homogeneity of the participants, only those with experience of living in the US were recruited. Certainly, future studies can investigate whether similar results can be obtained from Thai learners who have previously lived in other English-speaking countries. Moreover, at the time of the data collection, participants who had previous experience of living in the US had returned to Thailand. It was difficult to find and recruit only participants who had returned to Thailand for an exactly equal amount of time (e.g., one year). The TOEFL scores were used to ensure that the participants can be similarly characterized as being proficient enough in English to study in an English-speaking environment. The fact that the participants who graduated from a US university had returned to Thailand for no more than 3 years should not mean that, at the time of the current study, they were no longer capable of studying in such an environment, although arguably their English proficiency might have changed after return. However, to further reduce the possible effects of how recent learners were exposed to English in an ESL environment on their performances, future studies may include only learners who are still in an ESL environment at the time of data collection.

<sup>2</sup> An MI score is calculated from dividing the observed frequency of a word sequence in a specified span in a corpus by the corpus-based expected frequency and taking the logarithm to the base two of the result. A t-score is calculated from dividing the difference between the observed and the expected frequency of a sequence by the square root of the observed frequency (Gries, 2010).

<sup>3</sup> The mean score of each predictor was first subtracted from each observed value of that predictor. The result was then used to calculate the interaction term and used in the subsequent regression analyses.

## References

- Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the effects of exercises on verb–noun collocations. *Language Teaching Research*, 18, 54–74.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Cohen, J., Cohen, P., West, S.G., & Aiken, L. S. (2003). *Applied multiple regression/ correlation analysis for the behavioral sciences* (3 ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Cowie, A.P. (1998). *Phraseology: Theory, analysis, and applications*. Oxford: Oxford University Press.
- Crossley, S., & Salsbury, T. L. (2011). The development of lexical bundle accuracy and production in English second language speakers. *International Review of Applied Linguistics in Language Teaching*, 49, 1–26.
- Davies, M. (2013). Google scholar and COCA-academic: Two very different approaches to examining academic English. *Journal of English for Academic Purposes*, 12, 155–165.
- Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics in Language Teaching*, 47, 157–177.
- Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research*, 26, 163–188.
- Ellis, N.C. (2011). Frequency-based accounts of SLA. In S. Gass & A. Mackey (Eds.), *Handbook of second language acquisition* (pp.193–210). New York: Routledge/Taylor Francis.
- Ellis, N. C. (2013). Second language acquisition. In G. Trousdale & T. Hoffmann (Eds.), *Oxford handbook of construction grammar* (pp. 365–378). Oxford: Oxford University Press.
- Evert, S. (2007). Corpora and collocations. In A. Lüdeling & M. Kytö (Eds.), *Corpus linguistics: An international handbook* (pp.1212–1248). Berlin, Germany: Mouton de Gruyter.
- Farghal, M., & Obiedat, H. (1995). Collocations: A neglected variable in EFL. *International Review of Applied Linguistics*

- in Language Teaching*, 33, 315–332.
- Foster, P., Bolibaug, C., & Kotula, A. (2014). Knowledge of nativelike selections in a L2. *Studies in Second Language Acquisition*, 36, 101–132.
- Granger, S., & Bestgen, Y. (2014). The use of collocations by intermediate vs. advanced non-native writers: A bigram-based study. *International Review of Applied Linguistics in Language Teaching*, 52, 229–252
- Gries, S.T. (2010). Useful statistics for corpus linguistics. In A. Sánchez & M. Almela (Eds.), *A mosaic of corpus linguistics: Selected approaches* (pp. 269–291). Frankfurt, Germany: Peter Lang.
- Groom, N. (2009). Effects of second language immersion on second language collocational development. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language: Multiple interpretations* (pp. 21–33). Basingstoke, UK: Palgrave Macmillan.
- Prapaisit de Segovia, L., & Hardison, D. M. (2008). Implementing education reform: EFL teachers' perspectives. *ELT Journal*, 63, 154–162.
- 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: New perspectives on assessment and corpus analysis* (pp. 29–56). Amsterdam: John Benjamins.
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London: Routledge.
- Larsen-Freeman, D. (2003). *Teaching language: From grammar to grammaring*. Boston, MA: Heinle.
- Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, 29, 694–716.
- Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61, 647–672.

- Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18, 85–102.
- Li, M., & Kirby, J. R. (2014). The effects of vocabulary breadth and depth on English reading. *Applied Linguistics*, 36, 611–634.
- Li, P., Cadierno, T., & Eskildsen, S. W. (2014). Tracing an L2 learner's motion constructions over time: A usage-based classroom investigation. *Modern Language Journal*, 98, 612–628.
- Loewen, S., & Plonsky, L. (2016). *An A-Z of applied linguistics research methods*. New York: Palgrave Macmillan.
- McCarthy, M., & O'Dell, F. (2005). *English Collocation in Use – Intermediate*. Cambridge: Cambridge University Press.
- McCarthy, M., & O'Dell, F. (2008). *English Collocation in Use – Advanced*. Cambridge: Cambridge University Press.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp.35–53). Cambridge: Cambridge University Press.
- Meara, P. (2009). *Connected words: Word associations and second language lexical acquisition*. Amsterdam: John Benjamins.
- Muñoz, C. (2008). Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics*, 29, 578–596.
- Nassaji, H. (2006). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *Modern Language Journal*, 90, 387–401.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. New York: Cambridge University Press.
- Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: John Benjamins.
- Ortega, L. (2015). Second language learning explained? SLA across 10 contemporary theories. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed.) (pp. 245–272). New York: Routledge.
- Parkinson, J. (2015). Noun-noun collocations in learner writing.

- Journal of English for Specific Purposes*, 20, 103–113.
- Qian, D.D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, 56, 282–307.
- Qian, D.D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52, 513–536.
- Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10, 355–371.
- Read, J. (1998). Validating a test to measure depth of vocabulary knowledge. In A. Kunnan (Ed.), *Validation in language assessment* (pp. 41–60). Mahwah, NJ: Erlbaum.
- Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing*. Philadelphia, PA: John Benjamins.
- Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, 12, 328–363.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. London: Palgrave Macmillan.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64, 913–951.
- Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework. *Studies in Second Language Acquisition*, 19, 17–36.
- Schmitt, N., Ng, J. W. C., & Garras, J. (2011). The word associates format: Validation evidence. *Language Testing*, 28, 105–26.
- Schoonen, R., & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing*, 25, 211–236.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review*, 64, 429–458.

- Sonbul, S. (2015). Fatal mistake, awful mistake, or extreme mistake? Frequency effects on off-line/on-line collocational processing. *Bilingualism: Language and Cognition*, 18, 419–437.
- Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative methods. *Functions of Language*, 2, 1–33.
- Stæhr, L.S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, 31, 577–607.
- Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, 30, 79–95.
- Webb, S., & Kagimoto, E. (2009). The effects of vocabulary learning on collocation and meaning. *TESOL Quarterly*, 43, 55–77.
- Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing. *Studies in Second Language Acquisition*, 35, 451–482.
- Wolter, B., & Yamashita, J. (2018). Word frequency, collocational frequency, L1 congruency, and proficiency in L2 collocation processing. What accounts for L2 performance? *Studies in Second Language Acquisition*, 40, 395–416.
- 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.



**Appendix: Target Adjective-Noun Collocations**

<b>No.</b>	<b>Collocations</b>	<b>MI scores</b>	<b>t-scores</b>
1	simple reason	5.74	22.12
2	strong preference	6.63	10.24
3	big problem	4.79	39.66
4	strong opinion	3.74	7.33
5	heated debate	9.48	18.44
6	big decision	2.74	11.53
7	great importance	5.00	19.93
8	brief chat	6.49	4.31
9	large percentage	6.60	23.88
10	simple solution	6.54	15.88
11	harsh criticism	8.58	10.46
12	quick look	5.40	30.21
13	great detail	5.95	21.2
14	deciding factor	8.80	13.39
15	longstanding tradition	9.38	9.16
16	low wage	6.08	9.19
17	regular customer	6.47	8.84
18	strong wind	4.95	13.89
19	heavy rain	7.81	22.81
20	widespread damage	5.96	6.06
21	great sadness	5.54	9.83
22	absolute beginner	7.79	2.44
23	big meal	4.01	9.64
24	great pleasure	5.81	21.90
25	strong smell	4.60	8.29
26	large quantity	6.91	11.69
27	strong dislike	5.45	4.03
28	deep hatred	5.32	4.67