Factors Impacting Collocation Recognition and Controlled Production by ESL Speakers

by

Olga Makinina

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Applied Linguistics and Discourse Studies

Carleton University
Ottawa, Ontario

Abstract

The topic of collocations is popular in vocabulary acquisition-based studies. Researchers (Boers, Demecheleer, Coxhead, & Webb, 2014; Lewis, 2000; Pellicer-Sánchez, 2017; Webb, Newton, & Chang, 2013) implement many strategies for teaching/testing collocations, varying from intuitive noticing to repetition/drilling and analysis-based activities. However, collocational competence is slow to develop in learners of English as a second/additional language (L2), and so far, there is no clear understanding of factors impacting collocation acquisition. This dissertation examines external and internal factors influencing collocation recognition and controlled production by speakers of English as L2 in three consecutive studies.

The first study was based on an acceptability judgement-focused reading-comprehension task and a survey that compared populations of speakers of English as a first language (L1) and L2. This phase identified that speakers of English as L2 with higher test scores implemented similar reading-comprehension strategies as speakers of English as L1. Additionally, the factors of English as a predominant language of communication, L1, and vocabulary learning strategies were the most significant external factors. The results of the second and third stages of the research were based on a teaching module, pre- and post-instruction cloze and multiple-choice/acceptability judgement matrix-based tests, and a post-test survey. The findings suggested that such interconnected characteristics of collocations as frequency of occurrence and co-occurrence and morphosyntactic type (MST) influence collocation recognition and production. In terms of frequency, collocations with low frequency of occurrence and high frequency of co-occurrence were found to be more difficult to recognize pre-instruction and more susceptible to teaching

intervention. In terms of MST, Verb + Noun, Adjective + Noun, and Verb + Adjective were strong pre-instruction test score predictors; however, they were not easily acquired, and it was Verb + Adjective and Adverb + Adjective collocations that showed an improvement in post-instruction test scores. Other research implications concern strategies of teaching/testing collocations based on their characteristics. The testing format of a cloze and multiple-choice/acceptability judgment matrix modelled after Revier (2009) was an effective measurement of recognition and production of whole collocations in the sentence-level context. Additionally, a blended teaching approach based on noticing, repetition, and deep processing proved successful.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

ACL – Academic Collocation List (Ackermann & Chen, 2013)

ANOVA – Analysis of variance

BNC – British National Corpus

COCA – Corpus of Contemporary American English

CANCODE – Cambridge and Nottingham Corpus of Discourse in English

COBUILD – Collins Birmingham University International Language Database

COLLEX – Collocating lexis **(**Gyllstad, 2007, 2009)

COLLMATCH – Collocate matching **(**Gyllstad, 2007, 2009)

CONTRIX – Constituent matrix (Revier, 2009)

EAP – English as a second language for academic purposes

ESL – English as a second language

F – Frequency of occurrence

JACET – Japan Association of College English Teachers

L1 – First language

L2 – Second language

MI – Mutual Information

MICASE – Michigan Corpus of Academic Spoken English

MST – Morphosyntactic type

NS – Native speaker

NNS – Non-native speaker

PICAE – Pearson International Corpus of Academic English

P-P – Probability plot

PPMCC – Pearson product-moment correlation coefficient

SLA – Second language acquisition

Glossary

Collocation – A fixed semantically and structurally semi-restricted word combination of a minimum of two lexical words that are either adjacent to one another or separated from each other by 2-4 words.

Collocate (noun) **–** Dependent word(s) in collocation that are semantically and structurally subordinated to and regulated by a node. See Hausman (1989) and Sinclair, Jones, and Daley (2004).

Frequency of Occurrence – The number of times a word combination appears in corpora; it is usually measured by raw or normed frequency per million. See Halliday (1966), Korosadowicz-Struzynska (1980), and Siyanova-Chanturia and Spina (2015).

Frequency of Co-occurrence – The number of times words are used together in corpora; semantic and associative links between words; it is usually measured by mutual information. See Barnbrook, Mason, and Krishnamurthy (2013), Firth (1957), and Greenbaum (1970).

Morphosyntactic Type – Syntagmatic structural relations between words, which decide what parts of speech and word forms can be combined. See Benson, Benson, and Ilson (1986), Gitsaki (1996), Hausmann (1989), and Hill, Lewis, and Lewis (2000).

Node – The main word in collocation that is functionally, semantically, and structurally independent from its collocates. See Hausmann (1989) and Sinclair et al. (2004).

Span – The number of word components that precede or succeed the node in a collocation; the distance between collocation components. For example, in the

sentence "The boy took his classmate's place," the collocation is *to take [his classmate's] place*, and, therefore, the span is two words. See Firth (1957), Granger and Bestgen (2014), and Sinclair et al. (2004).

Chapter 1: Introduction

Overview of the Problem

Collocations, i.e., semi-fixed formulaic sequences with mutual expectancy between their components, are considered major operational units of language and as such, they have long been the focus of the second language (L2) vocabulary acquisition-related research. According to Gitsaki (1999), Handl (2008), Liao (2010), and Schmitt (2010), collocations perform a variety of functions, including conceptual (naming a concept/term), stylistic, and socio-pragmatic (function of social interaction for a specific communicative purpose). When denoting a concept, collocations help to express the meaning in a more concise way and thus contribute to speech efficiency and fluency (Hill, Lewis, & Lewis, 2000). In terms of style, as McCarthy and O'Dell (2005) suggest, "collocations give you alternative ways of saying something, which may be more colourful/expressive or more precise" (p. 6). For example, instead of repeating *very red*, a language learner might use the word combination *vivid red* or *flaming red*. As for the socio-pragmatic function, as explained by Henriksen and Stoehr (2009) and Wray (2002), collocations render the speaker's/writer's individual and social identity and facilitate social interaction by serving as codes that give certain clues about the communicative situation and the addresser's and addressee's social and professional status.

Consequently, collocational competence can be linked with communicative and socio-pragmatic competences and is closely related to the overall language proficiency development of English as a second language (ESL) speakers. Hargreaves (2000), Henriksen (2013), Nation and Webb, (2011), and Wang and Good (2007) suggest identifying several levels of collocational competence that can help teaching and

1

learning. These levels include the process or recognition of the form and meaning, i.e., ability to identify and single out necessary collocations in the oral and written language input; the understanding of semantic, morphosyntactic, distributional, and stylistic restrictions that regulate word choice in collocations; and the ability to use collocations productively in speech and writing. This usage is characterized by fluency, accuracy, and contextual/stylistic appropriateness.

At the same time, Hill (2000), Pellicer-Sánchez (2017), and Szudarski and Carter (2016) suggest that collocational competence is a domain of language knowledge that L2 learners most frequently lack, and while teachers often focus on correcting their grammar or sentence structure, it is faulty collocation use that becomes the source of most language errors. Over the decades, researchers and educators have been implementing different approaches to teaching and testing collocations, which range from implicit noticing, repetition, and acceptability judgement-based instruction discussed in recent studies by Pellicer-Sánchez (2017) and Peters (2014) to explicit lexical analysis-focused teaching discussed in studies by Boers, Demecheleer, Coxhead, and Webb (2014), Lackman (2011), and Lewis (2000). However, research by Bardovi-Harlig and Bastos (2011) and Nesselhauf (2003, 2005) suggests that even advanced L2 speakers with many years of prior instruction, both formal and immersion-based, still experience difficulties when recognizing and producing collocations.

Many studies have attempted to identify the reasons for this long-lasting challenge. Some researchers (Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2015; Yamashita & Jiang, 2010) draw parallels between L1 and L2 similarities and/or differences that might impact collocational knowledge. Other researchers compare L1

2

and L2 speakers' collocational knowledge and focus on extralinguistic factors of underdeveloped intuition (Siyanova, 2010; Siyanova-Chanturia & Martinez, 2014) and biological age of onset and cutoff of learning (Granena & Long, 2013). A different approach is to blame inadequate teaching/learning strategies, often targeted at early and teenage learners rather than adult L2 learners, for not addressing holistic processing of vocabulary units and offering insufficient language input (Durrant & Schmitt, 2010; Macis & Schmitt, 2017; Wray, 2008; Szudarski & Carter, 2016). A few studies dig directly into characteristics of collocations (Ellis, Simpson-Vlach, & Maynard, 2008; Nguyen & Webb, 2016; Peters, 2016) and examine the types of collocations that are likely to cause difficulties for L2 learners.

This last category of studies targeted at collocation-specific factors has been the least explored because despite over 60 years of phraseology and corpus-based research on collocations initiated by Firth's (1957) pioneering work on words and their associations, there is still no clarity in the definition and classification of this lexical phenomenon. The concept has traditionally been vaguely defined in shades of grey as semi-transparent, semi-restricted, and semi-substitutionable word combinations. Still, researchers cannot agree on the criteria; thus, collocations are often confused with other formulaic sequences. For example, Conklin and Schmitt (2008) and Tran (2011) consider collocations as types of idioms, whereas Gitsaki (1996) and Lewis (2000) include phrasal verbs as collocations.

A number of researchers (Biber, Conrad, & Cortes, 2003; Chen & Baker, 2010) suggest eliminating the concept of collocations altogether and focusing on lexical bundles, i.e., word combinations that frequently co-occur. However, others (Birch, 2014;

Heng, Kashiha, & Tang, 2014) insist that collocations are basic lexical units, while lexical bundles are considered as a sub-type of more frequent collocations as contrasted to less frequent collocations. Thus, collocation-targeted research is greatly in need of a clearer understanding and a comprehensive definition of the concept, as well as a classification of the factors that impact collocation acquisition. This is the aim of the present study: to identify the most and least useful collocations for learning ESL and thus contribute to a design of more effective testing and teaching tasks.

Objectives and Research Questions

This three-stage mixed-methods study is built around three research questions:

1. What external (extralinguistic and interlinguistic) and internal (collocation-specific) factors influence collocation acquisition by adult ESL learners?

2. How can internal (collocation-specific) factors be used for teaching and testing collocations?

3. What is the optimal format for testing recognition and controlled production of whole collocations?

Research Question 1 concerns identification and classification of the factors that could speed up or impede the collocation acquisition process in ESL learners. Although collocation acquisition is a life-long, largely intuition-based process in L2 (Siyanova-Chanturia & Martinez, 2014; Siyanova & Schmitt, 2008), existing research on collocations has identified a number of such factors, which can be classified into two domains. The first domain includes external non-collocation-specific factors: (1) extralinguistic, such as the age of onset and offset of learning English (Granena & Long, 2013; Han, 2004; Wray, 2002); length of residence in an English-speaking country

4

(Bardovi-Harlig & Bastos, 2011; Groom, 2009; Nesselhauf, 2005); language input and output (Adolphs & Durow, 2004; Durrant & Schmitt, 2010; Szudarski & Carter, 2016), and psychological-affective factors of culture and motivation (Li & Schmitt, 2010; Smith, 2005; Wray, 2008); and (2) interlinguistic, such as whether English is used as the predominant language for thinking and communication (Wang & Shih, 2011; Xu, 2015), English language proficiency (Groom, 2009; Laufer & Waldman, 2011; Nizonkiza, 2012, 2015; Schmitt, 2013), and interlanguage interference (Irujo, 1986; Liao, 2010; Millar, 2011; Peters, 2016; Wolter & Yamashita, 2015).

The second domain of potentially influential factors includes intralinguistic (or collocation-specific) factors, such as frequency of occurrence (Sonbul, 2015; Szudarski & Conklin, 2014; Wolter & Gyllstad, 2013) and co-occurrence (Durrant & Schmitt, 2009; Granger & Bestgen, 2014; Siyanova & Schmitt, 2008), morphosyntactic structure (Almela, 2011; Boers et al., 2014; Pellicer-Sánchez, 2017), and others.

Most research studies focus on one or two factors. To date, only a few studies (for example, Ellis et al., 2008; Macis & Schmitt, 2017; Ngyuen & Webb, 2016; Peters, 2016; Schmitt, Dörnyei, Adolphs, & Durow, 2004; Webb & Kagimoto, 2011) have attempted to compare and classify several factors in their relation to one another and collocation acquisition. Following these studies, my research objective is to examine, analyze, and compare these external and internal (collocation-specific) factors in their relation to one another and in terms of their statistical significance in order to comprise a comprehensive typology of factors that might impact collocation recognition and production.

Research Question 2 focuses on collocation-specific factors that can be used for selecting collocation items for testing and teaching and creating teaching and learning

5

tasks. At present, with the exception of a few studies by Nguyen and Webb (2016), Peters (2016), Revier (2009), and Webb and Kagimoto (2011), very little research explores how collocation characteristics impact collocation acquisition and might be implemented in teaching and testing. The majority of tests and learning exercises are based on the general principles of learning one-word vocabulary units, and collocation items are predominantly selected on the basis of frequency (Nesselhauf, 2005) or theme (Sarvari, Gukani, & Khomami, 2016; Schmitt & Schmitt, 2005). This method has not been quite effective; according to research by Granena and Long (2013) and Nesselhauf (2005), even after many years of instruction ESL learners still cannot adequately recognize and use collocations. Learning about basic collocation characteristics might help ESL learners master formulaic word combinations. Thus, the second research objective is to measure the effectiveness of teaching and testing collocations based on their frequency of occurrence in the corpora, frequency of co-occurrence (strength of association), and morphosyntactic type.

Finally, Research Question 3 is targeted at finding an optimal format for testing collocation recognition and production, which together comprise collocational competence. The majority of collocation tests in the above-mentioned studies are multiple-choice, fill-in-the-blank, translation, or matching exercises that do not measure knowledge of whole collocations, do not incorporate both recognition and production, and/or rely on the participants' L1. Consequently, the third research objective is to pilot and measure the effectiveness of such tests that are targeted at whole collocations, address both aspects of collocational competence (recognition and production), and at the

same time are manageable and standardized, i.e., measure controlled rather than free production of collocations.

Overview of Methodology

In order to address these three questions, this study implements a mixed methods approach that follows sequential transformative design (Creswell & Plano Clark, 2011). A mixed methods approach enables an examination of the issue from different angles, working with diverse populations of research participants, and triangulating data sources, collection, and analysis procedures (Yin, 2009). According to Creswell and Plano Clark (2011), a mixed methods approach might be implemented if one data source is not sufficient or if there is a necessity to come to generalizable conclusions; both of these are applicable to this study that seeks to combine and explain the relationships between different independent variables and the collocation acquisition outcome.

The study is divided into three closely connected consecutive stages (Stages I, II, and III) that focus on (1) exploring and developing a general picture of external factors that might be influential on collocation recognition; (2) identifying internal (collocation-specific) factors that impact collocation recognition and controlled production; and (3) developing and testing the effectiveness of a collocation-teaching module targeted at collocation-specific factors. The findings from each stage inform the subsequent stage(s); at the end, all findings are discussed in terms of a comprehensive typology of factors impacting collocation acquisition, as well as the testing and teaching tasks that address these factors. While large-scale SPSS-based quantitative data analysis prevails at all three stages, some findings are explained and supported by in-depth qualitative analysis (mostly drawn from participants' perspectives at Stage I).

7

Stage I: Methods overview. This study explored external extralinguistic and interlinguistic factors influencing recognition of false collocations by speakers of English as L1 and L2. For this purpose, it first implemented a pilot test that contained a multiple-choice task and an acceptability judgement task supplemented by post-test interviews with five participants. The interview data were subjected to qualitative structural and descriptive coding procedures after Saldaña (2012) in order to examine the major patterns that might indicate potentially significant factors. These factors served as a focus for the subsequent large-scale quantitative study with 93 participants (50 speakers of English as L1 and 43 speakers of English as L2). The acceptability judgement task in the test and test items were identical to the test used for the pilot study, and the post-test survey questions were based on the factors that emerged as significant during the pilot study interviews. These factors, coded as continuous (e.g., age or the number of years spent learning English) and categorical (e.g., gender, reading strategies, or vocabulary-learning strategies) independent variables, were analyzed using SPSS version 22.0 for their relation to the dependent variable of the mean test scores. First, these factors were analyzed independently of one another through a two-way between-groups and one-way analysis of variance (ANOVA) and independent samples t-test procedures. Second, correlation and multiple regression were implemented to explore the interaction of these factors and differences/similarities across the two groups of speakers of English as L1 and L2.

Stage II: Methods overview. This stage focused on collocation-specific factors influencing recognition and controlled production of whole collocations by ESL learners. The new test format, modelled after Revier (2009), was implemented where collocation

items were selected using a set of internal (collocation-specific) criteria. The post-test survey focused on the participants' language background and age in order to eliminate any possibly influential external factors. The data were coded as a set of independent interconnected variables. These variables represented five morphosyntactic types ranging on the scale of low - high normed frequency per million that measured the frequency of occurrence and low - high mutual information score that measured the frequency of co-occurrence, i.e., the strength of association. The data were analyzed by the means of SPSS. The predominant method of analysis was standard multiple linear, hierarchical, and stepwise regression that identified those collocation characteristics that were strong and weak score predictors.

Stage III: Methods overview. This final research stage tested the effectiveness of teaching based on collocation-specific factors and identified collocation types that were potentially problematic for teaching and learning. It employed a collocation-targeted teaching module and focused on the SPSS-based analysis of the pre- and post-instruction test results. The test items, format, and variables were identical to the test used at Stage II since those proved effective measurements of collocation recognition and controlled production. A series of paired samples t-tests were conducted to identify and compare statistically significant differences in the pre- and post-instruction overall mean test scores and mean scores in each category of five morphosyntactic types and frequency of occurrence and co-occurrence.

Informed Consent

The study was approved by Carleton University Ethics Review Board (Appendix A), where permission was granted to recruit participants, conduct tests and post-test

surveys, and conduct, audio-record, and transcribe interviews. At Stage I, five volunteer participants –one speaker of English as L1 and four speakers of English as L2– were recruited. Each participant signed a written consent form (Appendix B) that reserved their right to withdraw from the study up to one week after the test and interview. Next, 93 participants –speakers of English as L1 and L2– were recruited by the means of contacting their instructors through email and asking for their permission to administer a test and a post-test survey (Appendix B). These participants did not sign a written consent form because their participation in the test and post-test survey was anonymous to everyone, including their instructor. At Stage II, 78 participants were recruited in a similar way as at Stage I. At Stage III, participants were students in my own ESL classroom. Since their identities were not anonymous for me, an independent project team member obtained voluntary participants written consent (Appendix B). The data were accessed only after the final course grades were made available to the students, and the period for the final appeal of grades was over.

Organization of the Dissertation

This dissertation consists of nine chapters. Chapters 2-4 present a literature review of the current research on collocations. The focus of Chapter 2 is the definition and characteristics of collocations as outlined in the leading theoretical research approaches, namely phraseological and distributional approaches. Chapter 3 describes the findings of existing studies focused on the external and internal factors influencing collocation acquisition. Chapter 4 summarizes current testing formats and methodologies implemented for assessing collocation recognition, recall, and controlled production. Chapters 5-8 describe the findings of the three stages of the research. Chapter 5 focuses

on the external factors impacting recognition of false collocations. Chapter 6 presents the results of the research on the internal collocation-specific factors influencing collocation recognition and controlled production. Chapters 7 and 8 describe the teaching module and the findings of the pre- and post-instruction test. Chapter 9 provides a comprehensive overview of the study findings, discusses possible future directions, and limitations of the study.

Chapter 2: History of Collocational Studies and Characteristics of Collocation

Background

Although over the past decade collocation has become an increasingly popular topic in applied linguistics, the phenomenon remains notoriously hard to pin down because of the fuzziness of its characteristics and the obscure relations between collocational components. Even though the first reference to collocation as a placement of two and more items together can be found in the dictionaries going back as far as the 17th century (for example, *An English Expositor* by Bullokar [1616]) and collocation was defined as a syntactic construction in the reports of lexicographers at the beginning of the 20th century (for example, *Second Interim Report on English Collocations* by Palmer [1933]), it was not until Firth's (1957) famous "You shall know a word by the company it keeps" (p. 179), i.e., the meaning of a word is identified through its collocability (relations) with other words, that the concept of collocation began to be widely used in linguistics and education. Notwithstanding the long history of collocational studies, up until this day, the two opposing approaches to collocation studies –distributional (frequency-based) and phraseological (form and meaning-based)– and their varied combinations are surprisingly similar in the vagueness of their definitions.

Collocation in distributional studies. The distributional approach focuses on the statistical analysis of syntagmatic (i.e., morphosyntactic) relations (Mitchell, 1971; Nesselhauf, 2003, 2005; Sinclair, 1970, 1991, 2004) between the components of multiword combinations. This method of analysis involves extraction and counting of the statistically significant word chunks from corpora irrespective of their idiomaticity (Biber & Conrad, 1999, p. 990). The extracted formulaic sequences are then called

"collocational frameworks" (Renouf & Sinclair, 1991), "lexical bundles" (Biber, Conrad, & Cortes, 2003), or "n-grams" (Stubbs, 1995). Distributionalists often do not divide formulaic languages into sub-classes and, therefore, do not differentiate collocations from idioms or phrasal verbs, nor do they describe distinctive characteristics of collocations in terms of syntax or semantics. When analyzing formulaic sequences, the distributionalists' primary concern is the frequency of occurrence of the word combinations in corpora and the frequency of co-occurrence of the word components, i.e., the probability that these words are not encountered together by chance. These determine the statistical significance of a formulaic sequence.

The distributional approach is also called "corpus-based", where researchers in this field have created corpus-based collocational dictionaries, such as Collins Birmingham University International Language Database (COBUILD) English Language Dictionary (Sinclair, 2001); and word lists, such as Academic Word List (Coxhead, 2000), List of the Most Frequent Collocations in Spoken English (Shin & Nation, 2008), List of Frequent Collocations (Jaén, 2007), Academic Formulas List (Simpson-Vlach & Ellis, 2010), Phrasal Expression List (Martinez & Schmitt, 2012), and Academic Collocation List (Ackermann & Chen, 2013).

The definition prevailing in the distributional approach is that collocation is a frequent (statistically significant) co-occurrence of words at a certain distance from one another (Greenbaum 1970; Halliday, 1966; Sinclair, 1991; Sinclair, Jones, & Daley, 2004). This distance is identified as "collocational span". The criterion of frequency of occurrence prevails in all distributionalist classifications of collocations. For example,

Sinclair (1991) divides collocations into "regular", i.e., frequent, and "casual", i.e., non-frequent.

Collocations in phraseology. Characteristics of collocation are better addressed in the phraseological or "significance-oriented" (Herbst, 1996) approach, which is largely based on the linguistic analysis of the relations between the components in the word combinations. These specific relations bind words together and, therefore, can be called restrictions. Researchers (Benson, Benson, & Ilson, 1986; Cowie, 1998; Mel'čuk, 1998; Vinogradov, 1947, 1977) differentiate categories of multiword units that range from free non-bound word combinations (e.g., *play a game*) to fixed idioms and proverbs (e.g., *play the second fiddle*). Phraseologists attempt to identify and describe a set of standard criteria that characterize collocations (Cowie, 1998; Granger & Paquot, 2008), and so have developed a range of descriptive dictionaries –for example, English-Russian Phraseological Dictionary (Kunin, 1984), Oxford Dictionary of English Idioms (Cowie, Mackin, & McCaig, 1983), and BBI Dictionary of English Word Combinations (Benson et al., 1986; 1997).

The basic criteria in the field of semantics include the binary oppositions of semantic opaqueness versus transparency of meaning (Barfield & Gyllstad, 2009; Cowie, 1998) and fixedness versus variability of the relations between words (Nesselhauf, 2005), or, as Cowie (1998) and Barfield and Gyllstad (2009) put it, "substitutionability" or "compositionality". Non-compositionality of multiword units implies that the "global meaning [of the unit] is different from the sum of its individual parts" (Granger & Paquot, 2008, p. 31). The co-occurrence of highly phraseological formulaic sequences, such as idioms and proverbs, cannot be explained by the meaning of their individual

components, while the co-occurrence of free word combinations is easily explained and "predicted from the lexical meaning and semantic traits of a given lexeme" (Allerton, 1984, p. 30). Morphological criteria usually do not go beyond the restrictions as to what parts of speech can enter a formulaic word combination. In terms of syntax, phraseologists focus on the syntactic relations between components in restricted and free word combinations (Benson et al., 1986; Hausmann, 1989; Nesselhauf, 2005).

Based on these categories, the definition of collocation in phraseology moves beyond the frequency of occurrence and co-occurrence towards the description of semantic (associative), morphosyntactic, and pragmatic restrictions between collocation components as a result of similar lexical environments (Evert, 2008; Halliday & Hasan, 1976; Leech, 1974). In terms of semantics, collocations are viewed as combinations of words where the relations between them are fixed, yet at the same time allow some extent of variability, and the meaning of which is somewhat transparent (Nesselhauf, 2005; Vinogradov, 1977). Pragmatically, they represent formulas of speech used for different communicative purposes. Lexically, they are perceived as single units of meaning (Mel'čuk, 1998). In terms of syntax and morphology, Benson et al. (1986), Hausmann (1989), and Lewis (2000) provide classifications of lexical and/or grammatical collocations. In short, while grammatical collocations have a dominant lexeme (Noun, Adjective, or Verb) combined with a preposition or another grammatical construction, lexical collocations are combinations of two interdependent lexical words (Benson et al., 1986, p. ix).

It is possible to see that, as in the case of the distributional approach, the phraseological definition of collocations is still somewhat vague. For instance, to what

degree on the "phraseology continuum" (Granger & Menier, 2008) should a word combination be semantically transparent or restricted in its compositionality to be considered a collocation, free word combination, or an idiom? We can conclude that even multiple criteria are not enough to provide a comprehensive definition, since collocation is a fluid concept, and the boundaries between collocations and other formulaic sequences are often not clear or easy to explain. These boundaries are mostly reinforced through individualized usage (Granger & Paquot, 2008) and deducted through intuition (Wood, 2010); that is why the phraseological approach to collocation studies is often considered a top-down or deductive approach.

Characteristics of Collocations

When discussing collocations, some of the studies of the past decade (Ackermann & Chen, 2013; Schmitt, 2010; Wolter & Yamashita, 2015) tend to blend both approaches to provide a more comprehensive description. They initially select word combinations from the corpora based on their frequency of occurrence and co-occurrence and then analyze them using additional phraseological criteria. Following this blended taxonomy, this section of the chapter focuses on the semantic, morphosyntactic, frequency-based, pragmatic, and other characteristics that allow collocations to be classified as a lexical unit.

Transparency/opacity. According to Firth (1957), Sinclair et al. (2004), and Vinogradov (1977), collocational meaning is expressed through the whole word combination rather than through the sum of its individual components. This holistic meaning influences the degree of transparency/opacity, or "the degree to which words are used with their dictionary meaning" (Li & Schmitt, 2010, p. 25). According to Cowie

(1998), Handl (2008), Lewis (2000), Siyanova-Chanturia, Conklin, and Schmitt (2011), and Macis and Schmitt (2017), different collocations might be more or less semantically restricted and transparent depending on whether their individual components have a literal or figurative meaning in a particular context. For example, Aisenstadt (1979), Cowie (1998), Howarth (1996), and Nesselhauf (2005) classify Verb + Noun collocations into more or less semantically restricted based on the characteristics of the verb. Collocations with delexicalized verbs, such as *make, do*, and *have,* are less transparent than collocations with lexical verbs, e.g., *break into an apartment*, but more transparent than collocations with a verb that has a metaphorical meaning, e.g., *break into conversation*.

     Substitutionability (compositionality). This characteristic is directly related to semantic opacity. Since collocation is perceived as a meaningful whole, its components often cannot be replaced. According to Wanner, Bohnet, and Giereth (2006) and Zhu and Fellbaum (2015), decomposable collocations are usually those in which one word strongly depends on the other to the extent of blending in and losing its individual lexical meaning. Mel'čuk (1998) identifies four types of such collocations: (1) collocations with delexicalized verbs, such as *do a favor*; (2) collocations in which the meaning of the dependent word is fully clarified only through its relation with the main word, such as *black coffee*; (3) collocations in which one of the elements has a synonym, yet this synonym is impossible in a given word combination, such as *strong* (but not *powerful*) *coffee*; and (4) collocations in which the dependent word embraces the meaning of the main word, such as *aquiline nose* (p. 31). However, there are other collocation types that

17

allow for some variability, which brings us to the next collocational feature, strength of association.

Strength of association. This concept denotes the tendency of words to "co-occur in a certain semantic environment" (Hoey, 2005, p. 22). When words have stronger associations, they tend to co-occur more often and be less variable, i.e., decomposable, and more semantically opaque. McCarthy & O'Dell (2005) offer an example of such a collocation: *mitigating factors* (p. 8). Weaker associations imply a larger degree of compositionality and transparency. The words in such collocations are more independent from one another semantically. For example, the word *earnings* can collocate with several other words such as *annual*, *taxable*, *corporate*, all of which are fairly common collocations. Some researchers also differentiate the additional categories on the scale of the strength of association, such as "unique" or "fixed" collocations (McCarthy & O'Dell, 2005, p. 8) that do not allow for any changes (e.g., *back and forth*) and "medium-strength collocations" (Hill et al., 2000, p. 63; Revier, 2009) that are in the middle of the association strength continuum (e.g., *keep a dog*) and are essentially collocations with delexicalized and/or polysemantic verbs.

Morphosyntactic structure. Researchers suggest that collocations can be defined by the syntagmatic restrictions that decide what parts of speech and word forms can comprise collocations. While de Souza Hodne (2009) considers all collocational sets having a common lemma (different forms of the same word) as belonging to the same collocational range, Hoey (2005), Sinclair et al. (2004), and Stubbs (1995) argue that collocation analysis needs to be based on specific word forms rather than their lemmas, and as such, collocations of different forms of the same words will not be in the same

range. Moon (2011) adds to this that morphemes might impose certain restrictions on both collocational components and thus bind them. She gives an examples of Adjective + Noun collocations with the adjectival suffix *-like*, which in some cases contributes to figurative meaning of appearance, as in *jewel-like colors*, and in other cases intensifies literal meaning of shape, as in *star-like flower* (p. 498-503).

The majority of classifications of collocations (Benson et al., 1986; Gitsaki, 1996; Hausmann, 1989; Hill et al., 2000) are nevertheless based on their morphosyntactic types rather than morphemes. In these classifications, it is important to make a distinction between lexical types, which are formed by combinations of nouns, adjectives, verbs, and adverbs (for example, Verb + Noun, Adjective + Noun, Adverb + Adjective, Verb + Adverb, Verb + Adjective, or Noun + Noun), and grammatical types, which are comprised of a lexical part of speech and a grammatical component, such as a preposition or an infinitive clause. The latter are sometimes called colligations (Hoey, 2005; Mitchell, 1971) to differentiate them from collocations.

While some researchers focus on lexical collocations, claiming that only two and more meaningful words might form collocations (Hausmann, 1989), others include both lexical and grammatical collocations, thus blurring the border between collocations and phrasal verbs or combinations of nouns and prepositions (Benson et al., 1986; Hill, 2000). Additionally, Gitsaki (1996) and Lewis (2000) blend the criteria or morphosyntactic classes, word functions in a sentence, and semantics of the words. As a result, Gitsaki (1996) divides Verb + Noun collocations into two different types depending on whether the meaning of the phrase would be "creation" (for example, *make an impression*) or "eradication" (for example, *reject an appeal*) (p. 165), while Lewis

19

(2000) positions discourse markers, such as *put it another way*, and phrasal verbs as two separate sub-types of collocations (p. 133).

It is easy to see that such mixed classifications lead to either dividing those collocations that possess common characteristics or combining collocations with other formulaic sequences; therefore, for teaching and testing purposes, the classification of lexical collocations only is more convenient. Lewis (2000) and Sinclair et al. (2004) also connect collocations to the syntactic restrictions and differentiate between position-free (e.g., *terribly miss* or *miss terribly*) and position-dependent (e.g., *never mind* but not *mind never*) word combinations.

Node and collocate. The attempts at classifying collocations into different morphosyntactic types have been useful since they allow researchers to see collocations as combinations of a node (key word; base) and its collocates (values) (Hausmann, 1989; Mel'čuk, 1998; Sinclair et al., 2004). The concept of a node and its collocate(s) is related to the previously discussed semantic restrictions and strength of association. According to Handl (2008), if a node word can combine with many collocates, such collocations allow for more variation, are less restricted, and have weaker links between components.

Researchers differ in their understanding of relations between a node and its collocate(s). While Hausmann (1989) and Mel'čuk (1998) argue that collocations represent a hierarchical structure with the node being the leading element and collocate(s) as its dependents, Sinclair et al. (2004) suggest that both items are equal in their status. According to them, a node is "an item whose total pattern of co-occurrence with other words is under examination", while a collocate is "any one of those items which appears with the node" (p. 10). The existence of a node and collocate in collocation does not

necessarily mean that one collocational component might predict the other. The majority of modern corpus-based studies (Nguyen & Webb, 2016; Pellicer-Sánchez, 2017; Webb & Kagimoto, 2011) support Sinclair et al.'s (2004) view. For example, Webb and Kagimoto (2011) explore collocational ranges of nodes-adjectives in the Adjective + Noun collocations, and Nguyen and Webb (2016) focus on nodes-nouns with collocates-adjectives/verbs in the Adjective + Noun and Verb + Noun collocations.

Span. When analyzing collocational structure, researchers also discuss whether a node and its collocate need to be adjacent, and if not, what is the maximum span (the distance between the node and its collocate). While Firth (1957) does not specify an exact number of words and instead states that it might be anything from 2 to 15 words as long as it does not exceed a sentence, other researchers are more specific. Halliday (1966) considers a span of 3 words as sufficient, Sinclair et al. (2004) expand this limit to 4 words, and Gablasova, Brezina, and McEnery (2017) suggest 5 words. Despite the different views on the optimal span size, researchers generally agree that adjacent collocations have stronger associations between their components.

Frequency. The factor of frequency has traditionally been the cornerstone of the distributional approach as outlined in the works of Firth (1957), Halliday (1966), and Sinclair (1991, 2004). However, there is more than one dimension of frequency. The majority of researchers (Korosadowicz-Struzynska, 1980; Siyanova-Chanturia & Spina, 2015; Sonbul, 2015; de Souza Hodne, 2009; Szudarski & Conklin, 2014; Webb, Newton, & Chang, 2013; Wolter & Gyllstad, 2013) focus on the overall frequency of occurrence of word combinations in the corpora, which is measured by the raw or normal frequency per million. Different frequency thresholds have been established depending on the size

of the corpora, the length of the word string, collocational structure, and general study objectives. For example, Ackermann and Chen (2013) consider the normed frequency of one time per million in the Pearson International Corpus of Academic English (PICAE), which comprises 37 million words; Wolter and Gyllstad (2013) focus on the minimum threshold of 10 in the Corpus of Contemporary American English (COCA), which has 560 million words; and Szudarski and Conklin (2014) consider those collocations that occur more than 79 times in the British National Corpus (BNC), which has 100 million words.

To avoid inconsistency, Evert (2008) argues for the necessity of different approaches depending on the research objectives. While the first approach is based on a certain frequency threshold, the second approach uses ranking (i.e., word combinations are described as more and less frequent) with no minimum or maximum threshold, and the third approach combines both because it establishes the frequency threshold and ranks word combinations as more and less collocational. More recent studies indicate a preference for this latter strategy while establishing frequency thresholds and focus on frequency bands ranging from high-medium-low frequency collocations (Siyanova-Chanturia & Spina, 2015) to ten bands (Wolter & Gyllstad, 2013).

However, Sinclair et al. (2004) realize the insufficiency of analysis based on the frequency of the whole collocation and suggest classifying collocations according to the frequency of their individual components. They divide collocations into "upward", in which the collocate is more frequent than the node, for example, *very disappointed*; "downward", in which the node is more frequent, for example, *bitterly disappointed*; and "neutral", in which nodes and collocates have approximately the same frequency, for

example, *extremely disappointed* (p. xxiii). Sinclair et al. (2004) also mention that the strength of association between the components of upward collocations is statistically weaker and the components provide more information about their grammatical properties; however, the strength of association between the components of downward collocations is stronger and the components provide more information about the semantic properties of a word.

The frequency of co-occurrence is another frequency dimension addressed by the distributionalists. It is directly related to the strength of association between collocational components. It has traditionally been defined as the statistically significant probability that words co-occur together in the corpora (Firth, 1957; Greenbaum, 1970; Halliday, 1966; Sinclair et al., 2004) and is usually calculated using z-score, t-score, or mutual information (MI) score (Barnbrook, Mason, & Krishnamurthy, 2013). The MI score has been the most widely used measurement because it not only ensures that the words in a multiword unit co-occur on a regular basis, but also determines whether this co-occurrence is statistically significant, i.e., these words combined together have a certain shade of distinctive meaning and are thus strongly associated with one another to the point where one word can predict the other. In other words, the MI score determines if the word combination is formulaic. For example, although the word combination *social structure* is quite frequent (10.95 per million), it has a relatively low MI score (3.91) and is, therefore, less formulaic than such collocation as *stark contrast*, which is not particularly frequent (1.26 per million) but has a larger MI score (10.17).

As with the frequency of occurrence, there is a minimum threshold for the frequency of co-occurrence and strength of association. Most researchers agree that it

should be no lower than 3 (Hunston, 2002); however, as Evert (2008) states, 5-10 co-occurrences and more would make for stronger collocations. The frequency of co-occurrence stands as the basis of the two most widely spread approaches to collocation analysis. One of them is the analysis of lexical bundles (Biber et al., 2003) or n-grams (Stubbs, 1995) based on the extraction of two and more word combinations from the corpora irrespective of their idiomaticity. Another popular method served for the design of the COBUILD Dictionary and is based on extracting and counting statistically significant co-occurrences and differentiating formulaic sequences "through the sorting of the concordance lines by the first character after the keyword" (Sinclair et al., 2004, p. 46).

Length of collocational string. Researchers consider the number of words in a formulaic sequence as the determining feature of collocations. While Firth (1957) considers any word combinations of two and more words as collocations, Jones & Sinclair (1974) lower the number to a maximum of two non-consecutive words. Kjellmer (1994) further narrows it down to two adjacent words, which are defined by Vechtomova, Robertson, and Jones (2003) as "short span collocations", and by Sinclair et al. (2004) and Granger and Bestgen (2014) as "bigrams".

Difference between collocations and other formulaic sequences. Researchers supporting a phraseological approach to formulaic sequences have different opinions on the relationship between collocations and other formulaic units. Some researchers perceive collocations as sub-classes of idioms (Conklin & Schmitt, 2008; Tran, 2011) or "routine formulas" (Krishnamurthy, 2006), while others consider idioms and phrasal verbs as sub-types of collocations (Gitsaki, 1996; Hill et al., 2000; Poulsen, 2005).

Collocations and idioms. Researchers (Cowan, 1989; Fellbaum, 2007) determine the following similarities between collocations and idioms: (1) a multi-word combination; (2) non-compositionality of its components; and (3) morphosyntactic relations and restrictions between words in a phrase. However, they state that both collocations and idioms might have different degrees of these characteristics, with collocations being less morphosyntactically restricted and having some varied lexical components, for example, *hit the deck/sack/hay* (Fellbaum, 2007, p. 12). Cowan (1989), Fellbaum (2007), and Mounya (2010) also name semantic opacity as the main differentiating feature, with collocations being more semantically transparent.

Collocations and phrasal verbs. Though some linguists (Gitsaki, 1996; Lewis, 2000; Mounya, 2010) identify phrasal verbs as grammatical collocations, others (Hausmann, 1989; Hill et al., 2000) note that collocations need to have at least two meaningful lexical components in their structure. Therefore, Verb + Preposition + Noun structures might be considered collocations (e.g., *come down with the flu*), but Verb +Preposition or Particle structures are phrasal verbs (e.g., [the flag] *came down*).

Collocations and routine formulas (clichés). Barnbrook et al. (2013) state that collocation has a metaphorical meaning and is used in common situations, and can therefore be considered a cliché. However, Cowie (1998), Melc'uk (1998), and Nattinger and DeCarrico (1992) discuss one basic pragmatic function and semantic "hollowness" of clichés such as *Good luck!* and *How are you?* as contrasted with collocations that primarily carry out a certain semantic function, yet can often take on various pragmatic functions depending on the communicative context.

As follows from the literature, characteristics of collocation are fluid, multi-levelled, and interconnected. However, their combination leads to the following operational definition that will be used in this study. Collocation is defined as a combination of two lexical words (a node and a collocate) that frequently co-occur in the corpora, have a relatively strong mutual expectancy (measured by the MI score of 3+), and are either adjacent to one another or separated from each other by a span of 2-4 words. This word combination is characterized by different degrees of frequency of occurrence and co-occurrence in the corpora, transparency and figurativeness of meaning, and restrictions in compositionality (substitutionability), which is based on syntagmatic (morphosyntactic) relations of a node and its collocate.

Summary

This chapter situated the concept of collocation in the theoretical context of distributional and phraseological approaches to defining and classifying collocation at the level of semantics, morphosyntactic structure, and frequency. It was established that rather than having clearly defined characteristics that would differentiate it from other lexical units –such as idioms, phrasal verbs, and routine speech formulas– collocation can be described as an intermediate link between free word combinations and frozen expressions. It is characterized by the varying degrees of transparency of meaning, non-compositionality, as well as by the frequency of the whole unit, its individual components, and the frequency of co-occurrence (i.e., strength of association or mutual expectancy) between its parts. Both distributionalists and phraseologists cannot quite agree on the threshold for the frequency of occurrence and co-occurrence, nor can they decide how much variability in the meaning and structure of collocation is acceptable.

26

Consequently, a comprehensive mixed multi-levelled approach to defining and explaining features of collocation is preferred for the implementation of selecting teaching and testing collocational items. The next chapter will describe studies that have explored these collocational characteristics as well as other extralinguistic and interlinguistic factors that influence collocation acquisition.

Chapter 3: Factors Influencing Collocation Acquisition

The multiple factors impacting collocation recognition and/or production addressed in the existing research can be divided into several distinctive categories, namely: external, that is, factors that are not related to characteristics of collocation as a lexical unit; and internal, that is, collocation-specific factors. The latter can be further separated into extralinguistic and interlinguistic subcategories, each of which will be discussed in this section.

External Extralinguistic Factors

The factors of influence that are not language-specific vary widely, from biologically conditioned to cognitive and psychological-affective, as well as those related to the social and learning environment.

Intuition. Probably the most frequently mentioned and the most elusive factor impacting collocational competence is intuition. Because the links between collocation components are not easily explained or memorized, and the meaning of collocations is not always literal (Macis & Schmitt, 2017), identification and production of such formulaic sequences appears to be based on an "internal feeling" which is presumed to come naturally to speakers of L1 and has to be developed in speakers of L2 (Kjellmer, 1990; Yorio, 1989). Intuition is directly connected to the fluency of processing and production of formulaic sequences (Siyanova, 2010; Siyanova-Chanturia & Martinez, 2014). However, as many researchers (Macis & Schmitt, 2017; McGee, 2006; Siyanova-Chanturia & Spina, 2015; Wray, 2008) state, the concept of intuition is somewhat ambiguous and subjective. Their studies on L1 and L2 speakers' intuition show that intuition might differ even among L1 speakers: both L1 and L2 speakers tended to

provide different –and sometimes inaccurate– judgements about certain collocations when compared to data in corpora such as the BNC or COBUILD. Therefore, although well-developed intuition does influence collocation fluency, it cannot be over-relied on.

The age of onset and cutoff. Another factor is biological, where there is considered to be a crucial age of onset and cutoff for collocation acquisition. That is, there are certain age periods within which second language development is considered the most and least likely. These critical periods are widely explored in theories of maturation constraints and the Critical Period Hypothesis (Han, 2004; de Wit, 2007). Hyltenstam and Abrahamsson (2003) and Krashen, Long, and Scarcella (1979) identify that, in spite of the fact that adult learners might have more developed and varied learning strategies and better understanding of morphosyntactic structures, language acquisition occurs most naturally and effectively in young children with the cutoff age of 6-7.

Wray (2002, 2008) applies this knowledge to the acquisition process of formulaic language patterns. She suggests that younger learners might be more capable of mastering formulaic sequences precisely due to the fact that while adults' learning and communicative competence largely operates through analyzing the relations between the components of a formulaic word combination, children's mind works holistically –that is, they perceive collocation as a whole unit that can be used for certain communicative purposes and then further broken into parts or modified if necessary. In this sense, the adult learners' tendency to dissect and interpret moprhosyntactic constructions might be detrimental to their ability to perceive and produce formulaic sequences. Later studies more or less expand upon Wray's findings. For example, Foster, Bolibaugh, and Kotula (2014) focus on the receptive knowledge of lexical and morphosyntactic units and

identify a clear cutoff age of 12. Similarly, Granena and Long (2013) find that the ability to acquire collocations is at its highest from age 0-6 and decreases after the age of 9.

However, the notion of critical periods is not absolute. Foster-Cohen (2001) and Han (2004) state that depending on the certain aspects of language and learning environment, critical periods might differ. Granena and Long (2013) add that unlike other language domains like pronunciation or grammar, collocational competence can be gradually improved throughout life, which means that the cutoff is not absolute and there might be additional factors, such as explicit instruction or natural language aptitude, that facilitate acquisition (p. 336). Similarly, through interviews with adult L2 speakers who started learning the language between the ages of 7 and 17, Lahmann, Steinkrauss, and Schmid (2016) notice that the age of onset, unless it is combined with other factors, does not appear to be crucial when acquiring lexical units. However, the authors also note that these findings are somewhat questionable because an earlier age of onset (less than 7 years old) was not addressed in the study.

The length of residence. The factor most frequently mentioned alongside the age of onset/cutoff is the length of residence in the L2 environment. The general agreement is that as the time spent in an English-speaking country increases, so does collocational competence (Foster, 2009; Groom, 2009; Schmitt, 2000). Foster (2009) and Groom (2009) identify a positive growth in the output of idiomatic expressions by L2 learners who studied abroad as compared to those learners who were educated in their home country. However, they add that collocation acquisition is a slow process. As described by Lahmann et al. (2016), the length of residence should be a minimum of five years. Foster et al. (2014) claim that while shorter exposure periods might suffice for younger

learners, language immersion of 12 years and more will be more influential in the case of adult learners.

Conversely, Nesselhauf (2005) finds only a small correlation between the number of years learners spent in an English-speaking country and a decrease in collocational errors. Similarly, Adolphs and Durow (2004), Bardovi-Harlig and Bastos (2011), and Erman, Denke, Fant, and Forsberg Lundell (2015) point out that it is not enough for English language learners to be physically present in an English-speaking environment; they also need to engage in interaction with L1speakers of English in different socio-cultural settings. Similarly, Forsberg Lundell and Lindqvist (2014) discuss factors of motivation to learn or communicate using the target language, which, along with living in a target-language country, tend to bring better results in terms of collocation acquisition.

L2 input and output. The findings from the above-mentioned studies lead the discussion to the next relevant factor - frequency and quality of L2 input and output. Some researchers predominantly focus on input; for example, Durrant and Schmitt (2010), Macis and Schmitt (2017), Pellicer-Sánchez (2017), and Szudarski and Carter (2016) find positive results in collocational progress for learners who spend time reading and being frequently exposed to collocations. Others, such as Adolphs and Durow (2004) and Bardovi-Harlig and Bastos (2011), find that for international students living in host families mere exposure to language might not be enough for progress, especially when learners spend time in groups of other international students who speak their primary language. Bardovi-Harling and Bastos (2011) observe a positive influence, however, when there is "intensity of interaction" –that is, both input and output– with L1 speakers of English in different communicative situations outside of their regular learning

environment. This could include, for example, shopping, playing sports, partying, and so on. As a result of these interactions, learners recognize more frequently used formulaic sequences and use them more freely in their speech (p. 347).

Instruction. Although authentic L2 input and output are important, Kuo (2009) and Wray (2008) go so far as to suggest that L2 immersion without systematic instructional support might actually be detrimental for adult learners because they are not able to acquire formulaic sequences simply by listening and imitating L1 speakers. They explain that children mostly use bottom-up learning strategies, which means that they acquire whole word combinations as related to a particular context. In contrast, adults have background knowledge of communicative situations and therefore memorize words appropriate for each of these situations before they actually need them. In order to accumulate an extensive vocabulary, adults operate top-down strategies. Once they learn a word combination, they break it into smaller units instead of processing it as a whole based on associations between words. As a result, the links between words get lost, and adult learners resort to creativity, approximation, synonymy, or paraphrase to compensate for the deficiency of prefabricated formulas, which leads to collocational errors. Therefore, specific vocabulary learning and teaching strategies targeted at adult learners are needed (see more explanation of these strategies in Chapter 7).

Some researchers (Ghasemi, 2003; Lakshmanan, 2006; Wray, 2008; de Wit, 2007) suggest that the older learners are, the more attention needs to be paid to learning and teaching collocations explicitly –that is, directly– in a range of authentic contexts. Others (Webb et al., 2012) argue that "natural" learning (by reading/listening and repetition/imitating) is still effective because it develops adult learners' fluency and

ability to react quickly. Others still (Durrant & Schmitt, 2010; Szudarski & Carter, 2016) stand for a combination of both approaches. Nevertheless, the predominant majority is in agreement that formulaic sequences, such as idioms and collocations, should be prioritized over learning single word units since they are more difficult to master (Lewis, 2000; Peters, 2014) and yet crucial for language competency.

  Psychological-affective factors. Other external factors, which are frequently discussed in the context of language immersion and residence in an English-speaking country, as well as preferred learning strategies, are in the psychological-affective domain and are more or less related to the personality of an L2 speaker. In particular, L2 speakers might be able to achieve higher levels of collocational competence based on their positive attitudes towards and motivation to use the target language (de Wit, 2007; Han, 2004). Motivation and attitudes are partly related to learners' socio-cultural backgrounds and desire to adapt to the target language community and culture rather than keep to their own culture (Ghasemi, 2003; Tarone, 2006; Tollefson & Firn, 1983), and partly formed by their social relations with the target language speakers. Vigil and Oller (1976) and Tarone (2006) mention that while positive interaction between L2 and L1 speakers might facilitate the former's learning progress, negative experiences and feeling inadequate might affect the L2 learner's resistance to change, ability to notice and analyze forms in the L2 input, and reluctance to interact (Tarone, 2006, p. 160).

  As Wray (2002, 2008) suggests, the stakes are higher for adult L2 speakers than for children because adults play by more complex social rules and encounter a wider variety of communicative situations in which they need to demonstrate themselves as qualified experts. Thus, the majority of adult learners might avoid speaking in the target

language and/or use a limited range of formulaic sequences if they do not feel confident in their language skills. Similarly, Adolphs and Durow (2004), Gitsaki (1999), Groom (2009), Li and Schmitt (2010), and Smith (2005) notice that L2 speakers tend to over-rely on word combinations that they know. Yan (2010) further explains that L2 adult speakers might only partially understand certain expressions, and therefore be unwilling to use them.

External Interlinguistic Factors

English as the predominant language for communication and thinking. Discussion of the extralinguistic factors of length of residence and quality/frequency of the input and output suggests that those learners who predominantly use L2 for communication develop collocational competence more efficiently; consequently, those learners who use L2 for thinking might achieve even better results. Gabrýs-Barker (2006), Wang and Shih (2011), and Xu (2015) demonstrate that using English for private (i.e., verbalized speech for oneself), inner (i.e., silent speech for oneself), and spontaneous speech might positively impact learners' collocational competence in terms of collocation output. However, the measurements that researchers use for analyzing inner or spontaneous speech are somewhat limited, which leaves a need for further investigation into the relationship between this phenomenon and collocational knowledge.

In an attempt to address inner speech, Gabrýs-Barker (2006) implemented a translation task and analyzed participants' verbalized planning in both the L1 and the L2. However, since translation already involves using both languages, this approach might not provide an accurate measurement of inner speech in only the L2. Wang and Shih

(2011) used a self-report method, whereby participants were prompted to focus on and analyze their inner speech; however, as a measurement tool, self-report has a certain subjectivity. Despite the obvious limitations of self-report methods, Guerrero (2018) suggests that self-questionnaires and introspective verbalization of thinking might still provide useful insights regarding language learning and production, especially when combined with other research strategies that are focused on the participants' personal opinion and perception.

English language proficiency. The majority of researchers connect collocational competence to overall level of L2 proficiency and agree that collocational errors might gradually decrease but never disappear even at the most advanced stages, and that limitations exist for both receptive and productive knowledge (Gitsaki, 1996, 1999; Groom, 2009; Laufer & Waldman, 2011; Nizonkiza, 2012, 2015). However, the specific characteristics of collocational errors differ depending on proficiency level and the pattern is not straightforward.

In terms of collocation production, according to Schmitt (2013), these errors can be classified into: amount of use, accuracy/appropriacy of use, and quality/speed (p. 5). For amount of use errors, Eyckmans (2009), Granger and Bestgen (2014), and Nizonkiza (2015) notice that less advanced learners tend to overuse certain high frequency collocations, while stronger but low frequency collocations are mostly underused. Accuracy/appropriacy of use errors comprise the largest percentage of collocational errors and concern the ability of L2 speakers to use the correct lexical and morphosyntactic (grammatical) form of collocations. For lexical errors (e.g., paraphrasing, word-for-word translation, or blending different collocations), Laufer and

Waldman, (2011) and Wray (2002) suggest that more advanced learners actually make more mistakes because they tend to use formulaic sequences more freely and, not knowing exactly what constitutes appropriate collocations, often combine the wrong components. Otherwise, according to Nesselhauf (2005), learners might use correct collocations in the wrong context. She gives the example of an L2 speaker who uses the word combination *train one's muscles* instead of the more frequent and contextually appropriate verb *exercise*: "Thousands of women became members in sports clubs and trained their muscles" (p. 117). Similarly, morphosyntactic (grammatical) errors in which a wrong part of speech ("morphological synonymy") or preposition is used (Miyakoshi, 2009) can also occur, although less frequently. One typical example would be the word combination *drive license* instead of *driver's license*.

Researchers vary somewhat in their opinions on accuracy of use errors. Gitsaki (1996) and Yorio (1989) state that as the learners' language proficiency grows, they progressively make fewer grammatical errors. However, Bardovi-Harlig and Bastos (2011) note that learners tend to use formulaic sequences before they master their structure, which might lead to an increase of grammatical errors for intermediate learners in comparison to beginners. Even so, Bardovi-Harlig and Bastos (2011) come to the conclusion that lexical ("use") and grammatical ("well-formedness") proficiency even out in the case of advanced L2 speakers.

The third category, quality/speed of use errors, is connected to intuition and fluency. As per Hill (2000), "collocation allows us to think more quickly and communicate more efficiently" (p. 54). Hill further suggests that L1 speakers are fluent because they store a sufficient stock of lexical chunks in their memory that they can

extract from their mental lexicon whenever it is necessary. At the same time, according to other researchers (Siyanova & Schmitt, 2008; Schmitt, 2013; Wang & Good, 2007; Wray, 2002, 2008), L2 speakers are less fluent when producing collocations and have a less developed intuition regarding which word combinations are correct and appropriate. Because L2 learners focus on processing and producing individual words instead of multiword combinations, they spend more time trying to operate L2 units of communication; however, that might change over time with exposure to collocations and practice. The assumption is that as collocations become ingrained in memory, they are produced automatically based on need and communicative situation, which increases overall speech fluency.

As is evident in the above paragraphs, most studies that explore the connection between L2 proficiency and collocational competence focus on collocation production. There are, however, a few researchers (Laufer & Waldman, 2011; Nizonkiza, 2012, 2015; Talakoob & Koosha, 2017) who compared production and recognition and state that receptive knowledge usually outpaces production, particularly at the beginner and intermediate proficiency levels. Nizonkiza (2015) adds that receptive knowledge develops faster at the lower proficiency levels and then, as per Talakoob and Koosha's (2017) study findings, at a more advanced stage it slows down to a certain extent, at which time productive knowledge gains momentum.

L1 influence. Many research studies connect collocational errors, particularly accuracy/appropriacy of use errors, to the influence of learners' L1. This influence manifests itself in several ways. When learners do not remember an exact collocation in the L2, they might resort to the strategy of approximation, conditioned by L1-

programmed thinking (Han, 2004); i.e., they either directly translate collocations word-for-word from L1 to L2, which is not effective because collocations do not consist of separate words but rather of chunks, or provide an approximate translation in which one or several components correlate with L2 while the rest belong to L1 (Smith, 2005).

The impact of L1 might depend on whether collocation is congruent (i.e., similar in L1 and L2) or incongruent (i.e., different in L1 and L2). Studies on congruency and L1 influence usually focus on the recognition/processing of collocations and implement variations of acceptability judgement tasks (Peters, 2016; Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2015; Yamashita & Jiang, 2010) while reaction times and error rates are measured. The test items in these studies might be L1-specific (i.e., incorrect), L2-only-specific (i.e., incongruent), and congruent collocations. Carrol, Conklin, and Gyllstad (2016) also use a reading task and eye-tracking to measure the number and duration of fixations while reading. They identify that, notwithstanding proficiency level, for the majority of L2 speakers congruent collocations present less of a challenge, suggesting they are easier and faster to process and recall. Yamashita and Jiang (2010), however, observe that once L2 speakers learn an item, it is later recognized easily, regardless of in-/congruency. Similarly, Carrol et al. (2016) state that speakers' L1 seems to be more important at earlier stages of learning for both types of collocations, yet its influence is somewhat diminished in the case of more advanced learners.

Few studies focus on collocation production and use as related to L1 congruency. For example, Nesselhauf (2003, 2005) analyzed Verb + Noun collocations produced by non-native English speakers and came to the conclusion that congruent collocations were less challenging and contained fewer errors. However, other studies disagree and discuss

the hidden danger of congruent collocations. Irujo (1986) identifies two types of congruent formulaic sequences: those that are somewhat similar to L1, and are easily comprehended, yet equally easily confused, and those that are identical and are, therefore, the easiest to produce. In case of the former, a collocation's similarity with its L1 equivalent can actually be misleading and might cause negative L1 transfer, which results in "mother-tongue gridding" (Hasselgren, 1994, p. 238) or "false cognates" (Smith, 2005, p. 34). An example of such false cognate would be the word combination *plastic operation* instead of *plastic surgery*, created after its Russian equivalent *пластическая операция* [plasticheskaya operatsia].

Intralingusitic Collocation-specific Factors

While multiple interrelated external factors influence collocation acquisition in terms of its recognition and production to some extent, it is important to remember that they might manifest themselves differently depending on collocation types. Hence, a large portion of collocation-targeted studies focuses on specific characteristics of collocation as a formulaic lexical unit in English, as will be discussed in this section of the literature review.

Frequency of occurrence. According to the research (Durrant & Doherty, 2010; Granger & Bestgen, 2014; Siyanova-Chanturia & Spina, 2015; Szudarski & Conklin, 2014; Wolter & Gyllstad, 2013), higher frequency collocations are generally perceived and produced more efficiently and with less errors by the L2 speakers. Additionally, the ability to recognize and produce frequent collocations increases along with learners' language proficiency level and overall vocabulary knowledge. This shows that L2 learners are actually capable of holistically memorizing frequent word combinations. At

the same time, according to Siyanova-Chanturia and Spina's (2015) comparative study of L1 and L2 speakers' accuracy judgement of collocations, medium frequency collocations might present more difficulties not only for L2 speakers but also for L1 speakers, which leads researchers to conclude that the extremes (very frequent and non-frequent collocations) are more easily noticed.

Some researchers suggest paying more attention to the frequency of individual collocational components (collocates or nodes). Although there are studies (for example, Al-Thubaity & Bazeem, 2017) that focus on the collocate and/or node frequency as a tool for extracting collocations from the corpora, so far only a few studies focus on the frequency of collocational components as an influential factor in terms of acquisition, namely collocation recognition. Hoffmann and Lehmann (2000), for example, compare English L1 and L2 speakers' receptive knowledge and find that collocations with less frequent collocates are recognized more easily.

Nizonkiza (2015) also studies collocation recognition, but focuses on the frequency of nodes. She chooses items for the collocation test from Nation's (2001) word list with the range of frequency bands from 2,000 to 5,000 times. Similarly, Nguyen and Webb (2016) examine L2 learners' receptive knowledge of collocations with the node words belonging to the 1,000 to 3,000 word frequency bands. Both of these studies show that downward collocations with more frequent nodes are easier to recognize for L2 learners. According to the researchers, one possible explanation for this might be that previous teaching and learning strategies focused on single frequent words and words they could combine with. Additional evidence of this supposition can be found in the study by Macis and Schmitt (2017) where, although they do not identify a statistically

significant correlation between the overall frequency of a word combination in the corpora and collocational knowledge, they suggest that the frequency of the individual words comprising a collocation might be a factor influencing overall collocation knowledge.

Frequency of co-occurrence. Although the frequency of occurrence of word combinations in a corpus might be a strong indicator of how easily a collocation can be recognized/produced, it is the frequency of co-occurrence of collocational components that helps to differentiate collocations from random, non-formulaic word combinations (Ellis et al., 2008; Evert, 2008; Nguyen & Webb, 2016; de Souza Hodne, 2009).

Studies by Durrant and Schmitt (2009), Granger and Bestgen (2014), and Siyanova and Schmitt (2008) indicate that frequency of occurrence and co-occurrence are interconnected rather than separate entities because they essentially measure the same thing: formulaicity of a word combination, i.e., how strong the link is between the components in a word combination and how often one might encounter this word combination. Thus, when reporting their research findings, Siyanova and Schmitt (2008) analyze typical collocations in the L1 and L2 speakers' corpus on the basis of the "joint frequency/MI criteria" (p. 437). Durrant and Schmitt (2009) and Granger and Bestgen (2014) likewise divide collocations into several categories based on raw frequency (measured by the t-score) and MI bands (from low to high).

Yet other studies by Ellis et al. (2008), Nguyen and Webb (2016), and Simpson-Vlach and Ellis (2010) focus on the differences between the frequency of occurrence (measured by the raw/normed frequency per million) and co-occurrence (the strength of association measured by the MI score) and attempt to compare them as two separate and

potentially influential independent factors that might impact collocation acquisition. This comparison yields some interesting results. For example, when comparing the processing of formulaic sequences by speakers of English as L1 and L2, Edmonds and Gudmestad (2014) and Ellis et al. (2008) discover that for speakers of English as L1, the strength of association impacts the speed of recognition/production. In other words, the higher the MI score, the faster participants process/produce word combinations. On the contrary, for speakers of English as L2, their recognition/production is driven by the frequency of occurrence of the word combinations.

Similarly, Nguyen and Webb (2016) show that collocation recognition by L2 speakers is impacted by the frequency of word combinations in the corpora and the node frequency, but not by their strength of association, that is, the frequency of co-occurrence. Additionally, Nguyen and Webb identify that collocations with higher MI scores, i.e., stronger collocations, present more challenges precisely because they are less frequent and/or have less frequent node words. These findings suggest that it is necessary to teach collocations that would not only be frequent in the corpora, but also have a high strength of association. A similar rationale led Simpson-Vlach and Ellis (2010) to select collocations with high raw frequency and MI score for their list of academic formulaic sequences. Similarly, Ackermann and Chen (2013) use normed frequency, t-score, and MI score as criteria for their Academic Collocation List.

Morphosyntactic structure. The node word and collocates that are associated with one another enter into syntagmatic relations, conditioned by their morphosyntactic characteristics. The majority of research focuses on one particular morphosyntactic type of collocation, most frequently either Verb + Noun or Adjective + Noun collocations. For

example, studies by Boers et al. (2014), Eyckmans (2009), Koya (2005), Nesselhauf (2005), Revier (2009), and Szudarski (2016) examine Verb + Noun collocations since these are traditionally considered the most widely used and problematic type. The results of these studies suggest that L2 speakers might experience difficulties when using collocations with delexicalized verbs, such as *do*, *make*, or *have*, as well as frequent verbs, such as *take*, *give*, and *hold*. Varied verb forms in terms of tense, modality, and voice present an additional challenge.

Studies by Jaén (2007), Pellicer-Sánchez (2017), and Siyanova and Schmitt (2008) explore Adjective + Noun collocations that prevail in the corpora and find that this morphosyntactic type is easier for acquisition because there is less variability – that is, adjectives and nouns do not change their forms and possess a more distinct lexical meaning. Nevertheless, it might be problematic for L2 learners to select the correct adjective due to the large number of synonyms that can be interchangeable in certain contexts and yet non-substitutional in collocations. For example, according to Siyanova and Schmitt (2008), the adjectives *ill* and *negative* are synonyms for the adjective *bad*. However, while the collocations *bad atmosphere* and *negative atmosphere* are fairly common in English, the word combination *ill atmosphere* is not used.

Yet fewer studies address other morphosyntactic types, such as Adverb + Adjective (Erman, 2014; Granger, 1998) or Verb + Adjective (Almela, 2011). These studies report that the difficulties learners experience when recognizing/producing such collocations are directly related to the variability of forms and the lexical meaning of collocational components. Thus, Almela (2011) examines "collocations of transitive verbs and attributive adjectives (e.g., *cause by* + *faulty/defective* + *Noun*)" (p. 49) and

identifies that such Verb + Adjective collocations represent a compound of Adjective + Noun and Verb + Noun collocations, and as such, would possess characteristics of both types.

Some researchers (Gitsaki, 1996, 1999; Handl, 2008; Mahvelati & Mukundan, 2012; de Souza Hodne, 2009; Szudarski & Carter, 2016; Walker, 2008) choose to incorporate different morphosyntactic types. However, the purpose of using them is to select a variety of testing/learning collocational items rather than to examine the impact of the morphosyntactic structure on collocation acquisition. Thus, Szudarski and Carter (2016) use Adjective + Noun and Verb + Noun collocations as items for their learning exercises that targeted input enhancement and post-instruction test. Gitsaki (1996) describes how collocational knowledge changes depending on the learners' proficiency level and, for this purpose, uses over 30 different morphosyntactic collocational constructions in her translation, blank-filling, and other writing exercises.

Only a few recent studies by Nguyen and Webb (2016) and Peters (2016) compare different morphosyntactic types as pertaining to collocation acquisition. Nguyen and Webb (2016) discuss recognition of Adjective + Noun and Verb + Noun collocations and find participants equally lacking in the receptive knowledge of these two types. Conversely, Peters (2016) finds some differences between these two types and confirms that Adjective + Noun collocations present less of a challenge than Verb + Noun collocations. According to Peters, this finding can be explained by less variation of adjective forms in comparison to verb forms (voice, tense, and modality) and a stronger association between the components of Adjective + Noun collocations versus Verb + Noun collocations, which helps L2 learners have whole collocations rather than their

separate components stored in memory. On the contrary, Wolter and Yamashita (2015) suggest that when compared to non-formulaic equivalent sequences, Verb + Noun collocations might be processed faster than Adjective + Noun collocations. At the same time, the authors admit that since participants in their study shared a common L1 (Japanese) and collocational items were selected based on the principle of their congruency/non-congruency, the reliability of this contradictory finding might be in question.

Other structural and semantic characteristics. While the frequency of occurrence and co-occurrence and the node-collocate relationship are the basic characteristics of collocations and, therefore, the factors most commonly addressed in the studies, some minor collocation characteristics also need to be mentioned. Research has found a correlation between L2 speakers' ability to recognize/produce collocations and structural characteristics, such as the length of collocational strings, collocational span, i.e., adjacency of lexical items comprising a collocation, and the number of collocates for the node word; and semantic characteristics, such as collocations of near-synonyms and transparency/opaqueness.

Length of collocational strings. After Sinclair et al. (2004) introduced bigrams (i.e., formulaic sequences consisting of two words) as opposed to n-grams (i.e., formulaic sequences of more than two words in length), researchers' opinions on the acceptable length of collocational strings have differed. Those studies that consider only lexical (notional) units comprising a collocation focus on binary constructions, that is, collocations consisting of two lexical elements (Almela, Cantos, & Sanchéz, 2013; Durrant & Schmitt, 2010; Granger & Bestgen, 2014). However, those researchers who

include both grammatical and lexical words compare collocations consisting of three to five elements (Simpson-Vlach & Ellis, 2010). The latter study's findings confirm that the longer collocational string, the more time it takes for L2 speakers to process it. For example, the two-words-long collocation *wide variety* would be memorized and processed faster than the multiword unit *circumstances in which it has been shown*.

      Span. The concept of length is directly connected to the notion of collocational span, i.e., the distance between collocational components, which is normally 2-4 words as introduced by Halliday (1966) and popularized by Sinclair et al. (2004). These studies indicate that a longer span leads to more variation and flexibility and less fixedness for word combinations, which means collocations with a longer span are less strongly associated. They are considered non-adjacent and, as research by Gyllstad (2007) and Vilkaité (2016) shows, they are processed –and, consequently, acquired– more slowly than their adjacent counterparts. For example, the collocation *provide information* would be read and processed faster than the collocation *provide some of the information*. At the same time, Vilkaité (2016) uses an eye-tracking experiment and identifies that non-adjacent collocations with three words in between the collocates, such as *provide some of the information*, are still processed faster than non-fixed word combinations of the same size, such as *compare some of the information*.

      Position of the node and number of collocates. Webb and Kagimoto (2011) examine adjacent collocations with different positions of the node word: before the collocate (e.g., *good laugh, reason,* or *behavior*) or after the collocate (e.g., *difficult, full,* or *right time*). They do not identify any significant relationship between L2 speakers' collocational knowledge and the position of the node. However, this study finds that the

number of collocates for a node has a positive influence on L2 speakers' collocation processing and learning. Specifically, learners were more successful in acquiring collocations in a set of the same node word and six different collocates (for example, *black mark/eye/humor/sheep/cloud/market*) versus three collocates (for example, *big business/deal/day*).

 Synonyms. A number of studies (Chan & Liou, 2005; Erman, 2014; Jafarpour, Hashemian, & Alipour, 2013; Webb & Kagimoto, 2011) examine collocations of near synonyms, i.e., words that are similar in meaning and yet enter different collocations (Xiao & McEnery, 2006), for example, *weak tea* and *feeble voice*. Researchers agree that collocations of synonymous words usually present an additional challenge for L2 speakers; unlike L1 speakers, L2 speakers do not possess strong mental associative links of word combinations and thus might interchange synonymous components incorrectly. Additionally, as Erman (2014) suggests, even those collocations that allow for interchangeable synonyms –that is, Adverb + Adjective collocations that might take different synonymous adverbs, such as *fully, completely*, or *entirely*– are still restricted and the links between them guide the speakers' choice of formulaic sequence.

 Transparency/opaqueness of meaning. Researchers (Gyllstad, 2007, 2009; Gyllstad & Wolter, 2016; Howarth, 1996; Jaén, 2007; Nesselhauf, 2005; Revier, 2009) have come to somewhat different results when studying the differences in L2 speakers' processing of more and less semantically transparent collocations. Most of them agree that opaqueness contributes to holistic perception, which is not developed in the case of L2 speakers, and is therefore a challenge. For instance, Siyanova-Chanturia, Conklin, and Schmitt (2011) conducted an eye-tracking experiment with speakers of English as L1 and

L2 and compared their processing of formulaic sequences with literal (i.e., direct) meaning, for example, *at the end of the day* meaning "in the evening", and figurative (i.e., indirect, metaphorical) meaning, for example, *at the end of the day* meaning "eventually". While in the case of speakers of English as L1 they find no difference between processing of these two types, speakers of English as L2 process figurative sequences more slowly because they need to re-read them. The extent of this problem requires further examination.

Howarth (1996), Nesselhauf (2005), and Wolter and Gyllstad (2013) suggest that semi-transparent collocations, in which one component has a literal meaning and another component has a figurative meaning (for example, *catch a cold* or *catch a glimpse*, in which verbs have a figurative meaning), are more difficult for L2 speakers than wholly opaque idioms (for example, *catch someone red-handed*) or free word combinations (for example, *catch a ball*). The reason for this can be found in Nesselhauf's (2005) and Wray's (2008) explanation: idioms are more likely to be perceived as "frozen" decomposable wholes with no separate meanings for individual components, while collocations presuppose a certain degree of composability and variation, while still being restricted. Macis and Schmitt (2017) add that in the latter case, speakers of English as L2 are more likely to perceive word combinations as individual components rather than the whole and might not be able to identify the figurative meaning clearly. This finding is slightly undermined by the fact that although Macis and Schmitt (2017) identify challenges that L2 speakers experience with semi-transparent collocations and point out that participants were able to recognize only 33% of collocations with figurative meanings, they do not find any statistically significant difference between the L2

speakers' judgement of the more and less semantically transparent collocations. When presented with a set of collocations and asked to rate them on the scale of transparency/opaqueness based on whether their meaning is difficult to guess, participants gave different ratings to collocations of similar transparency. For example, more participants selected the collocation *queen bee* as being easier to guess than the collocation *rainy day* (p. 333). The researchers explain this by the subjectivity of the L2 speakers' intuitions. However, another possible reason might be that, as per Cieślicka (2006) and Wray (2002, 2008), participants did not process either type of collocations holistically but focused on their individual lexical components instead.

Comparative Studies of Multiple Factors

Of the studies of the factors influencing collocation acquisition, only a few attempt to combine and compare different factors of influence. In this section, some of these studies will be discussed.

Comparative studies of multiple external (extralinguistic and/or interlinguistic) factors. Schmitt et al. (2004) examine such extralinguistic factors as age, gender, language aptitude, and motivation. They created a series of cloze and multiple-choice tests designed to measure learners' receptive and productive knowledge of formulaic language. These tests together with the language aptitude and motivation tests were administered to ESL learners. Results show that even in a 2-3-months-long course participants demonstrated improvement; however, that might be accounted for by the fact that participants had a more advanced proficiency level and demonstrated knowledge and awareness of formulaic sequences on the pre-test. Another explanation might the test format, where although participants had to complete a fill-in-the-gap task, some letters of

the missing formulaic sequences had already been provided, which might have speeded up the recognition process. Meanwhile, the factors of motivation and aptitude did not emerge as significant.

Another study by Bardovi-Harlig and Bastos (2011) discusses how proficiency level, the length of residence in an English-speaking country, and frequency and quality of the L2 input and use outside of classroom impact recognition and production of formulaic sequences by speakers of English as L1 and L2. The researchers administered a recognition task, which asked participants to listen to a set of formulaic sequences and identify whether they had heard these phrases before, as well as a production task, which required participants to deliver an impromptu speech on a given topic using formulaic sequences. The results indicate that length of residence does not have an impact on either recognition or production of formulaic sequences. The authors point out that most participants had resided in the country for eight months or less, which might not have been enough. On the other hand, proficiency and frequency and quality of the input prove statistically significant, showing that formulaic language competence increases with time if sufficient input has been provided.

Yamashita and Jiang (2010) address the factors of L1 congruency and frequency and the quality of L2 input as influential in terms of collocation recognition by speakers of English as L1 and L2. They used an accuracy judgement task, in which participants had to evaluate collocations from the 3,000 word frequency levels from the Japan Association of College English Teachers (JACET) List of 8,000 Basic Words (JACET Committee of Basic Words Revision, 2003) and pseudo-collocations while their reaction times and error rates were measured. The results demonstrate that there is no difference

in the processing of congruent and incongruent collocations by speakers of English as L1; yet, for speakers of English as L2, L1 incongruency is an important mitigating factor. However, the study also shows that with sufficient exposure to collocations, the impact of L1 congruency diminishes.

Carrol et al.'s (2016) study yields similar results. The researchers examine the factors of L1 influence, exposure to L2, and proficiency level as related to the processing of formulaic sequences. Speakers of English as L1 and L2 were asked to read sentences containing Verb + Noun formulaic sequences, which were selected on the basis of their congruency with the ESL participants' L1. Eye-tracking and analysis of reading time, fixation count, and skipping showed that L1 incongruency is a significant factor for speakers of English as L2; yet more advanced participants had less difficulty with English-only formulaic sequences, which means that proficiency and exposure to L2 might positively influence processing skills.

Comparative studies of multiple internal (intralinguistic, collocation-specific) factors. The research by Ellis et al. (2008) summarizes the results of several studies targeted at the frequency of occurrence, mutual information, and collocational length as processing of collocations by speakers of English as L1 and L2. Collocations with the varying span (3-5), frequency, and MI bands were the focus of three different tasks, namely identifying English-like word strings, reading collocations aloud while the voice onset and articulation time were recorded and analyzed, and reading aloud only the last word of formulaic sequences. The results show that although advanced speakers of English as L2 approximate speakers of English as L1 in terms of processing collocations,

significant differences are still present. While L1 speakers are influenced by mutual information, L2 speakers process collocations based on their frequency of occurrence.

Similarly, Nguyen and Webb (2016) examine how collocation frequency in the corpora, node word frequency, mutual information, L1 congruency, and morphosyntactic structure (Adjective + Noun and Verb + Noun) influence collocation recognition by speakers of English as L2. For this purpose, the authors selected bigrams where the node words belong to the 1,000-3,000 word frequency levels from the BNC and COCA for the recognition-based test, which required participants to select an appropriate collocate for a node word. The results indicate that the frequency of a node has a larger impact on collocation recognition than the frequency of a whole collocation due to the fact that collocations are generally less frequent than individual words and that L2 speakers might memorize collocates for the most frequent words only. L1 congruency and the MI score are also significant; yet, no evidence has been found for the significance of the morphosyntactic structure.

Peters (2016) conducts a similar study that also addresses L1 congruency, morphosyntactic structure (Adjective + Noun and Verb + Noun), and length of collocational strings as potentially influential factors. For the collocational items, the node words were selected from the high frequency lists in COBUILD English Dictionary for Advanced Learners (Sinclair, 2001), with the overall length of collocational strings ranging between 9 and 17 letters. Participants were provided with a list of collocations alongside fill-in-the-gap and find a synonym/antonym/hyponym tasks, and took a pre- and post-treatment translation-based test designed to measure their collocation recall/recognition. Like Nguyen and Webb (2016), Peters (2016) identifies the

significance of L1 congruency; yet, contrary to Nguyen and Webb, she does identify that morphosyntactic structure is a significant factor of influence, as well as the word length and the overall size of the learners' vocabulary.

The study by Webb and Kagimoto (2011) is somewhat different from the previous studies because it focuses on both structure and semantics as influencing productive collocational knowledge, and it examines the independent variables of the number of collocates, the position of the node word, and near-synonymous collocational sets. The participants had to memorize a set of Adjective + Noun collocations, which were non-congruent and adjacent and included frequent words from the first and second BNC word lists, and took a pre- and post-treatment translation-based test targeted at measuring productive collocational knowledge. The results show the significance of the number of collocates and synonyms. However, as the authors mention, the translation task might have impacted the participants' production of collocations, and it might have been better to provide them with definitions in the L2 instead.

Comparative studies of multiple external and internal factors. Nesselhauf (2003, 2005) focuses on both receptive and productive knowledge of Verb + Noun collocations and addresses a variety of factors, ranging from extralinguistic (such as the frequency and quality of the input and the age and length of exposure, and length of residence) to interlinguistic (such as L1 congruency) and intralinguistic (such as collocational restrictions in terms of the morphosyntactic structure and semantic transparency). She finds that the length of residence and formal teaching are of little significance, which means that being exposed to collocations is not enough, nor are currently used teaching strategies necessarily effective. Compared to external factors,

internal collocation-specific factors are more significant. L1 influence/congruency emerges as the most important factor since it contributes to collocational errors of transfer and paraphrasing. Semantic transparency also plays a role, with semi-transparent collocations being the most difficult for L2 learners due to their variability and, therefore, weaker mental links between the components of such collocations.

Similarly to Nesselhauf (2003, 2005), Koya (2005) explores the development of receptive and productive collocational knowledge. One of her research stages focuses on the internal factors of collocational restrictions, semantic transparency, structure (delexicalized/lexical verbs, prepositions, and articles), and meaning of individual collocational components, as well as the interlinguistic factor of L1 congruency. She identifies that in terms of productive knowledge, the degree of semantic transparency is the most significant factor, with congruency, structural simplicity, lexical verbs, and context-free meanings of collocational components being close forerunners, and collocational restrictions being the least significant. These results are explained in the study by the fact that the factors of semantic transparency, verb lexicalization, and meanings of individual components are interrelated, as are the factors of L1 congruency and structure. However, the findings are different for receptive knowledge; while semantic transparency no longer seems to have an impact, the factors of congruency, structure, semantics of individual components, and restrictions are all moderately influential.

The study by Szudarski and Conklin (2014) explores how the extralinguistic factor of input, the interlinguistic factor of L1 congruency, and collocation-specific factors of frequency and morphosyntactic type impact collocation processing by L2

speakers. For this purpose, the researchers selected Adjective + Noun and Verb + Noun collocations with a high MI score and divided them according to low-high frequency of occurrence. Participants had a practice/treatment session focused on rote rehearsal and took a pre- and post-treatment test, which involved reading L1-L2 sentences containing collocations in question and determining whether these collocations were English-like or not. Their reaction times and error rates were measured throughout the activity. The results correspond to other studies discussed in the previous section: morphosyntactic type, collocational frequency, L1 congruency, and frequency of exposure to collocations (i.e., quantity of the input) are all significant.

A more recent study by Macis and Schmitt (2017) examines the impact of internal (semantic transparency and frequency of occurrence) and external (time spent at a university in an English-speaking country and L2 communication outside of classroom) factors that impact figurative collocation recognition by speakers of English as L2. Participants performed reading comprehension tasks and provided definitions for collocations. The test items were Adjective + Noun and Verb + Noun collocations with different frequency ranges. The results suggest that the external factors of communication in L2 and the length of residence and studies in an English-speaking country do have a moderate impact; however, the internal factors of frequency and semantic transparency are not influential. This result contradicts Szudarski and Conklin's findings (2014), as well other studies summarized in the previous subsection. Macis and Schmitt offer several explanations. According to them, collocations that are frequent in the corpora might not be frequent in the learners' input. In addition, the frequency of whole collocations is lower than the frequency of individual words comprising collocations. As

for the factor of semantic transparency, researchers suggest that L2 speakers might not have a well-developed intuition that would help them transition between perceiving collocational components literally and figuratively.

Summary

This chapter discussed the factors previously identified in the research studies as potentially influential in terms of collocation recognition and production. The majority of studies focus on the external extralinguistic (intuition, the age of onset and cutoff, length of residence in an L2-speaking country, frequency and quality of input, learning/teaching strategies, culture, motivation, and attitudes) and interlinguistic (using L2 for communication and thinking, L2 proficiency level, L1 influence/congruency) factors, with fewer studies paying attention to the features of collocation, such as the frequency of occurrence and co-occurrence (measured by the strength of association), the frequency of a node and collocate(s), morphosyntactic structure, the length of collocational strings, collocational span, the number of collocates, synonymy, and semantic transparency. The small number of studies focusing on characteristics of collocation might be explained by the fuzzy nature of collocations, which, according to representatives of the phraseological approach, is positioned in between free word combinations and figurative idioms and allows for variation. Thus, it is hard for researchers to pinpoint the extent of influence of semantic and structural characteristics. It is also notable that only a limited number of studies address and compare multiple factors as related to one another and collocation acquisition; therefore, the overall picture of what impacts collocational competence remains unclear. The next chapter will explain research methods implemented in some of the most influential studies on collocation recognition and production.

Chapter 4: Methodologies for Collocation Acquisition Research

Corpus-based Versus Psycholinguistic Studies

The two most prominent approaches to researching the factors influencing collocation acquisition are corpus linguistics-based studies and psycholinguistic studies. The former explores characteristics of collocations in written (Groom, 2009; Laufer & Waldman, 2011; Li & Schmitt, 2010; Reppen, 2009; Siyanova & Schmitt, 2008) or oral corpora (Wang & Shih, 2011) of speakers of English as L2 and/or L1, or in ESL/EFL textbook corpora (Jiang, 2009; Koya, 2006). Conversely, psycholinguistic studies use testing and instructional tasks (Bardovi-Harlig & Bastos, 2011; Cieślicka, 2004; Conklin & Schmitt, 2008; Gitsaki, 1996; Jiang & Nekrasova, 2007; Peters, 2009; Wray, 2008) and laboratory experiments, such as brain scan and eye tracking while reading (Carrol et al., 2016; Wolter & Gyllstad, 2011, 2013; Yamashita & Jiang, 2010), to examine English as L2 speakers' collocational competence.

Corpus research is mainly used to identify formulaic sequences in oral and written speech (Conklin & Schmitt, 2008), explain their characteristics, and prepare data for subsequent psycholinguistic research (Weinert, 2010). Corpus linguistics mainly focuses on frequency of occurrence and co-occurrence counts that help identify recurrent word combinations and distributional patterns of collocational components. Often, a computer search draws researchers' attention to the formulaic sequences that, according to Wray (2002), "may 'intuitively' not seem formulaic, i.e., with no corresponding meaning/function attached", and, therefore, are not considered in the studies (p. 6). It is evident that while corpus-based studies might help to describe intralinguistic features of

collocations in the L1/L2 writing/speech, they cannot tell much about the external

processes and factors influencing collocation acquisition.

Psycholinguistics is said to step in when corpus studies are not sufficient. While

corpus studies provide an answer to the question of "what?", i.e., what are the features of

collocations in English as L1/L2 speakers' corpora, psycholinguistics answers the

question "why?", i.e., why are some collocations recognized/produced more easily, while

others are more challenging? As Weinert (2010) notes, psycholinguistic research serves

the purpose of interpreting and verifying corpus analysis data. Ellis et al. (2008) add that

psycholinguistic analysis might help to find out how language speakers perceive and

produce formulaic sequences, and what their difficulties are conditioned by. It taps into

some of the extralinguistic factors, such as reading fluency and span (Carrol et al., 2016;

Conklin & Schmitt, 2008), vocabulary learning strategies (Szudarski & Carter, 2016;

Wray, 2008), and interlinguistic factors, such as using L2 for communication and

thinking (Wang and Shih, 2011) and L1-L2 interference (Peters, 2016; Wolter &

Yamashita, 2015; Yamashita & Jiang, 2010).

Testing Collocation Recognition, Recall, and Production

Based on the objectives of the present research, which is focused on the overall

picture of factors influencing collocation acquisition, this section surveys selected

psycholinguistic and mixed-methods studies that include testing tasks and select

collocational items from the general corpora of English, such as the BNC or COCA.

Corpus-based studies that focus on learners' oral or written corpora and specialized

psycholinguistic experiments based on eye-tracking and brain scanning are not included

in this review since these methods are not applicable to the methodology used in this research as outlined in Chapter 1.

The studies that develop collocational tests can be roughly divided into three categories, namely recognition-, recall-, and production-focused tests. While production is usually defined as output, recognition is subtler and less straightforward. When explaining the recognition process, Siyanova and Schmitt (2008) note that it has several phases that include "(a) reading [and listening comprehension], (b) locating in the mental lexicon, and (c) making decisions" (p. 49). Similarly, Henriksen (2013) discusses "the initial process of recognition of the collocational unit, the process of mapping meaning or function on to form, expansion of knowledge of use restrictions of the unit, and the development of receptive... fluency" (p. 44). Recall, meanwhile, is yet another process and can be considered as a bridge between production and recognition. Gyllstad (2009) specifically differentiates between recall, in which "the form or the meaning of a word is retrieved and supplied when triggered by some sort of prompt stimulus", and recognition, i.e., identifying the correct word from different options (p. 156). Recognition and recall are easier to control for when developing testing experiments and tasks. As a result, testing free productive collocational knowledge is problematic because it involves a lot of variation and possible data outliers –that is why even those studies by Revier (2009) and Talakoob and Koosha (2017) that pose themselves as targeted at collocation production are in fact measuring controlled production/recall.

Although the list of the studies presented in this survey is selective rather than exhaustive, it highlights the major research on testing collocations during the past decade (2007-2017). Older studies by Gitsaki (1996), Koya (2005), and Schmitt et al. (2004)

have also been included because they present some of the earliest and most comprehensive attempts to explore internal and external factors impacting collocation acquisition.

As Table 1 shows, the most frequently used format for recognition tests is based on choosing an item, whether in a multiple-choice test or an acceptability judgment task, while the tasks of matching or providing definitions are less common. The choice of format is informed by the mental processes considered to be engaged in recognition. As per Gyllstad (2009), recognition involves noticing and identifying the correct word among other options, which is why providing choices seems to be an optimal task for targeting this process.

The principles that researchers use when selecting distractors for the test items differ. For Gyllstad (2007, 2009), Pellicer-Sánchez (2017), Siyanova and Schmitt (2008), Yamashita and Jiang (2010), and partially Nizonkiza (2015), the criterion for distractors is the low frequency with which these word combinations and/or their individual components occur in corpora. Other studies by Szudarski and Conklin (2014) and Wolter and Gyllstad (2011, 2013) focus on both the frequency of occurrence and MI score as well as implement non-existing "pseudowords" or "fillers". Two recent studies by Macis and Schmitt (2017) and Nguyen and Webb (2016) prioritize the MI score. For them, the main difference between collocations and random word combinations with distractors would be higher or lower MI scores, i.e., the frequency of co-occurrence. Finally, Nizonkiza (2015) brings in semantics and makes her participants choose between synonyms, i.e., participants are prompted to check the context of use as well.

Table 1

Chronological Overview of Recognition-targeted Tests

Studies	Methodologies	Test Items	Participants
Gyllstad (2007, 2009)	1. Collocating lexis (COLLEX) multiple-choice test between acceptable and non-acceptable collocations; 2. Collocate matching (COLLMATCH) grid-formatted test where participants chose between a formulaic and non-formulaic sequence	High frequency (5 or more per million) Verb + Noun collocations where nodes were retrieved from the JACET 8000 word list; and random free word combinations	Three groups: two groups of high school and university ESL learners and native speakers of English (NSs)
Siyanova and Schmitt (2008)	Acceptability judgment task in which participants evaluated collocations based on how common they are in English; reaction times and error rates were measured	Adjective + Noun collocations ranging on the scale of frequency of occurrence (from "typical" English-like to "atypical" non-English-like), retrieved from the BNC and checked against the BBI Dictionary of English Word Combinations (Benson et al., 1997) and the Oxford	NSs and advanced ESL/EFL learners from different L1 backgrounds, majoring in English and specializing in English as teachers or translators; divided into three groups based on length of residence in an English-speaking country

Studies	Methodologies	Test Items	Participants
		Collocations Dictionary for Students of English (Crowther, Dignen, & Lea, 2002)	
Yamashita and Jiang (2010)	Acceptability judgment task in which participants evaluated collocations based on how common they are in English; fixation points, reaction times, and error rates were measured	Adjective + Noun (span = +/-2) and Verb + Noun (span = +/-1) congruent and incongruent collocations with approximately the same frequency and length, retrieved from the BNC and Shogakukan Corpus Network; and non-English-like word combinations where node words were retrieved from the most frequent words in the JACET List	NSs, advanced ESL speakers (teachers, researchers, or students in the United States), and EFL learners who had never lived in an English-speaking country and did not have exposure to English outside of the classroom; both ESL and EFL speakers had approximately the same amount of formal language instruction and their L1 was Japanese
Wolter and Gyllstad (2011)	1. Lexical decision task in which participants evaluated collocations based on how common they are in English; reaction times and error rates were measured; 2. COLLMATCH test for ESL	Congruent and incongruent Verb + Noun collocations ranging in frequency (the minimum threshold was 5 per million), but having approximately the same length and frequencies of their individual	NSs and advanced ESL speakers whose L1 was Swedish

Studies	Methodologies	Test Items	Participants
	speakers only	components, checked against the Bank of English Corpus and the BNC; and word combinations with real verbs but pseudo-collocates ("fillers")	
Wolter and Gyllstad (2013)	Acceptability judgment task in which participants were presented with the first collocational component (the prime) and then identified whether the following word was real or not; reaction times and error rates were measured	Adjective + Noun congruent and incongruent collocations ranging in frequency (the minimum threshold was 10 per million), retrieved from the COCA; and non-English-like word combinations with "fillers"	NSs and advanced ESL speakers whose L1 was Swedish
Szudarski and Conklin (2014)	1. Pre-test and post-test with an acceptability judgement task in which participants decided whether a word combination was acceptable; reaction times and error rates were measured; 2. Treatment based on repetition	Adjective + Noun and Verb + Noun collocations ranging in frequency (threshold 30 per million), but similar in length and with high MI scores, retrieved from the BNC; and random word combinations ("fillers")	Advanced ESL speakers whose L1 was Polish; proficiency level was measured by Vocabulary Levels Test (Schmitt, Schmitt, & Clapham, 2001)

Studies	Methodologies	Test Items	Participants
	and underlining of collocations		
Wolter and Yamashita (2015)	Double lexical decision task in which participants were shown two words simultaneously with one word above the other and they decided if both words are English-like; reaction times and error rates were measured	Adjective + Noun and Noun + Verb incongruent English collocations, retrieved from the COCA on the basis of their MI score of 3 and more; English word-for-word translations of Japanese-only collocations; and random non-collocational word combinations where the prime was a pseudoword and the second component was a real word in English	NSs and two groups of intermediate and advanced ESL speakers whose L1 was Japanese
Nizonkiza (2015)	Multiple-choice test where participants selected a correct collocate; distractors were synonyms of collocates retrieved from dictionary.com	Verb + Noun collocations where node words were retrieved from Nation's (2006) 2,000-5,000-word bands and collocates were retrieved from the Oxford Collocations Dictionary for Students of English (Crowther et al., 2002), where the frequency of collocates was	ESL university-level learners whose L1 was Kirundi and Swahili; proficiency level was measured by the TOEFL test

Studies	Methodologies	Test Items	Participants
		approximately the same as the node frequency	
Nguyen and Webb (2016)	1. Multiple-choice test where participants selected a collocate for a node out of four options; distractors were selected among the words belonging to the same part of speech and with approximately the same frequency, yet with low MI score; 2. Matching test where participants matched sets of nodes and collocates	Verb + Noun and Adjective + Noun adjacent incongruent, partially congruent, and wholly congruent collocations selected from 1,000-3,000 word frequency levels from Nation's (2012) lists based on the BNC and COCA, where the frequency threshold was 50 times per million in the COCA, and where nodes and collocates (both real and distractors) were of approximately the same frequency level (the MI score for the node word and collocates was 3 and more, and for the node words and distractors, less than 1)	Pre-intermediate to upper-intermediate ESL learners (based on the scores of the university entrance examination) whose L1 was Vietnamese
Pellicer-Sánchez	1. Reading task in which participants read a story with	Adjective-pseudoword collocations, where the pseudowords-nodes	ESL learners from different language backgrounds; intermediate

Studies	Methodologies	Test Items	Participants
(2017)	target collocations containing "pseudowords"; these collocations were repeated 8 times; 2. Post-reading multiple-choice tests where participants provided the form and meaning of pseudowords and whole collocations and recognized the correct collocate for the pseudo-node; participants indicated their degree of un-/certainty about their choice	substituted the 1,000-3,000 high frequency nouns from the BNC and the adjectives were selected among the 20 frequent collocates of these nouns and from the 2,000 frequency level lists from the BNC	proficiency level (based on a self-assessment of their language skills)
Macis & Schmitt (2017)	1. Participants wrote a definition for collocations that were highlighted in bold in sentences; 2. Post-test questionnaire with questions about participants' background and the frequency/quality of English language input/output	Semi-restricted adjacent Adjective + Noun and Verb + Noun collocations retrieved from the COCA on the basis of their MI score of 3 and more	ESL speakers whose L1 was Spanish and who had either never lived in an English-speaking country or lived there for less than 20 months; no proficiency measurement

Another feature of research on collocation recognition is the principle of comparison. Gyllstad (2007, 2009), Siyanova and Schmitt (2008), Wolter and Gyllstad (2011, 2013), Wolter and Yamashita (2015), and Yamashita and Jiang (2010) compare speakers of English as L1 and L2 based on the speed and accuracy of their recognition of different types of collocations ranging in their frequency and L1 congruency. Meanwhile, Nizonkiza (2015), Macis and Schmitt (2017), Nguyen and Webb (2016), Pellicer-Sánchez (2017), and Szudarski and Conklin (2014) use ESL speakers only and focus on comparing different types of collocations that the speakers recognize and/or produce. As a rule, the latter studies address a wider range of collocational features, such as the MI scores, adjacency, and semantic and structural restrictions.

Table 2 demonstrates that recall-focused tests are typically implemented to demonstrate the effectiveness of different types of collocation-targeted instruction and include gap-filling exercises. The difference between this task and tasks measuring controlled production (as outlined in Table 3) is that the stimulus is provided in the form of the initial letters (Durrant & Schmitt, 2010) or mnemonic alliterations (Boers et al., 2013).

Table 3 focuses on controlled production-focused studies only because the tasks used in the free production-targeted studies, such as comparing samples of participants' writing (Groom, 2009; Laufer & Waldman, 2011; Li & Schmitt, 2010; Reppen, 2009; Siyanova & Schmitt, 2008) or oral elicitation experiments that collect and analyze participants' oral speech samples (Wang & Shih, 2011), are not quite feasible for testing, which pre-supposes a set of controlled conditions. As can be seen from Table 3, tests of

controlled production typically involve translation from L1 to L2 and include additional

clues, such as sentences in which these collocations can be used.

As Table 4 shows, the majority of studies on collocation acquisition use mixed

methods; that is, at different stages they use different testing tasks designed to measure

production, recognition, and/or recall since these processes are interrelated and together

constitute collocational competence. Mixed studies allow researchers to address a number

of factors impacting collocation acquisition and implement more varied teaching

treatments. Thus, these studies explore different populations of participants who range in

language proficiency levels, prior instruction, and exposure to English, as well as

different collocation characteristics, such as frequency of occurrence and co-occurrence,

frequency of nodes and/or collocates, congruency, semantic transparency, structural

restrictions, and alliterations, and different types of teaching intervention, such as

meaning- and form-focused, noticing while reading, listening, or writing, and repetitions.

Table 2

Chronological Overview of Recall-targeted Tests

Studies	Methodologies	Test items (units of analysis)	Participants
Durrant and Schmitt (2010)	1. Collocation-targeted instruction with three different foci: (a) single exposure to the items; (b) oral repetition of the same sentence; (c) repeated encounter with collocations in various contexts; 2. Fill-in-the-gap test which required participants to provide a node for an existing collocate based on the first two letters only	Low frequency Adjective + Noun collocations retrieved from the BNC with the high frequency of their components (nodes and collocates) of 50-100 per million; noun nodes are 4-5 letters-long	High intermediate ESL speakers from different language backgrounds; proficiency level was measured by IELTS and TOEFL tests
Boers, Eyckmans, and Lindstromberg (2013)	1. Treatment in which participants completed a dictation and identified collocations with repeating consonants; 2. Post-test in which participants listened to the dictation of single words from collocations encountered during the treatment session and provided another word	Alliterative and non-alliterative collocations of relatively high frequency (4,500 and more times in the COCA)	EFL learners whose L1 was Dutch; proficiency level was higher intermediate based on the minimum entry score for the enrollment in an English class

Table 3

Chronological Overview of Controlled-Production-targeted Tests

Studies	Methodologies	Test items (units of analysis)	Participants
Webb and Kagimoto (2011)	1. Pre- and post-instruction test where participants translated L1 collocations into L2; 2. Instructional treatment based on three-minute-long memorization of collocations, their translations, and sentences in which these collocations were used	Adjective + Noun adjacent collocations from 1,000-2,000 frequency level lists from the BNC; there were several collocational sets based on the following criteria: (1) L1 congruency; (2) position of a node before or after the collocate; (3) number of collocates (several collocates for each node word versus one node-one collocate sets); and (4) collocates-synonyms	ESL speakers whose L1 was Japanese, with prior instruction for six or more years; proficiency level was measured by a productive-knowledge-based pre-test
Kim (2017)	1. Pre- and post-treatment fill-in-the-blank tests where participants had to complete the sentence by providing the translation of an L1 collocation into L2;	Adjective + Noun, Verb + Noun, and Preposition + Noun collocations retrieved from McCarthy and O'Dell's (2005) textbook	EFL learners whose L1s were Chinese, Japanese, and Korean, of different levels of English language proficiency, participants had no prior experience of living/studying in an English-speaking country; proficiency level was

Studies	Methodologies	Test items (units of analysis)	Participants
	2. Collocation-noticing/analysis-targeted treatment; 3. Questionnaire on participants' demographic information and language background		measured by the entry test participants had to take in order to enroll in English courses

Table 4

Chronological Overview of Mixed Tests

Studies	Methodologies	Test Items	Participants
Gitsaki (1996)	1. Instruction based on exercises from collocation-targeted textbooks; 2. Free production task (essay writing) on one of the topics in the textbook; 3. Translation from L1 to L2; 4. Fill-in-the-blank task in which participants had to complete sentences by providing collocates; 5. Questionnaire on the participants' demographics and language background	Collocations belonging to different morphosyntactic types as per an expanded version of Benson et al.'s (1986) classification	Three groups of EFL learners whose L1 was Greek, ranging in proficiency level from less to more advanced; proficiency level was measured by the class level
Schmitt, Dörnyei, Adolphs, and Durow (2004)	1. Pre- and post-test: (a) Cloze test with fill-in-the-blank sentences where the initial letters of each word of the formulaic sequence was provided; (b) multiple-choice test with fill-in-the-blank sentences where multiple-choice	Formulaic sequences retrieved from (a) The Longman Grammar of Spoken and Written English (Biber et al., 1999) and Nattinger and DeCarrico's (1992) phrases, and their frequency examined with the	Advanced ESL learners whose L1 was mostly Chinese and Japanese; proficiency level was measured by TOEFL test scores

Studies	Methodologies	Test Items	Participants
	options were provided, where distractors were appropriate in English and yet not suitable in the given context; 2. Treatment with frequent and repeated exposure to and analysis of formulaic sequences; 3. Questionnaire on participants' attitudes towards language learning	BNC, CANCODE Corpus of Spoken English, and Michigan Corpus of Academic Spoken English (MICASE); and (b) seven textbooks used in the ESL classes	
Koya (2005)	1. Productive collocational test with fill-in-the-gap sentences written in English where participants translated collocations from L1 to L2 with the nodes already provided; 2. Receptive multiple-choice-based collocational test where participants identified a correct verb, where distractors were synonyms of the verbs	Verb + Noun collocations selected based on the criteria of different degrees of restriction, opacity, and L1 congruency, individual meanings of verbs and nouns, as well as their structure (presence of delexicalized verbs, prepositions, and/or articles)	EFL learners whose L1 was Japanese, with no length of residence in an English-speaking country and a minimum of six years of formal instruction; proficiency level was measured by the vocabulary size test developed after Mochizuki (1998) and Nation (1990)

Studies	Methodologies	Test Items	Participants
Jaén (2007)	1. Receptive collocational multiple-choice-based test where participants identified a correct collocation for a dictionary definition; 2. Productive collocational gap-filling test where participants completed a dictionary definition by providing an adjective-collocate for the noun-node	Semantically transparent Adjective + Noun collocations with the span of +/- 4; noun nodes retrieved from the most frequent words in the ADELEX list that comprises the Bank of English, BNC, and Longman Corpus Network corpora; collocates selected from among the most frequent co-occurrences in the BNC	ESL university students majoring in English linguistics whose L1 was Spanish
Eyckmans (2009)	Discriminating Collocations Test (DISCO) based on the multiple-choice format where participants chose between collocations and pseudo-collocations	Verb + Noun semi-restricted collocations ranging in frequency; verbs selected from the General Service List (West, 1953) and nouns selected from the BNC and COBUILD Bank of English	ESL upper-intermediate learners; proficiency level was measured by the L1-L2 oral proficiency task
Revier (2009)	1. Constituent matrix (CONTRIX) test with a sentence with a cloze gap and a range of choices for the node and collocate in the grid, where distractors in the grid were all appropriate	Verb + Noun adjacent collocations with polysemous verbs, such as *get, make, raise,* and so on, ranging on the scale of semantic transparency and frequency (from .04 to .47 per	EFL learners whose L1 was Danish, with different proficiency levels; proficiency level was measured based on the high

Studies	Methodologies	Test Items	Participants
	collocations in English; however, only one word combination would fit into the context of the sentence; 2. Questionnaire on participants' demographics and English language learning strategies	million in the BNC); noun collocates selected from the 3,000 most frequent words in the Web VocabProfile (Cobb, n.d.); sentences retrieved from PIE/GOOGLE concordances	school or university class they attended
Bardovi-Harlig and Bastos (2011)	1. Recognition task in which participants listened to and identified word combinations they have heard often/not; 2. Production task in which participants produced spontaneous speech on a given topic; 3. Questionnaire on participants' language, length of residence, and time spent interacting with NSs and non-native speakers (NNSs) of English	Formulaic sequences of two and more words ranging on the frequency scale	NSs and low-intermediate to low-advanced NNSs of English
Webb, Newton, and Chang	1. Multiple-choice-based pre-test where participants selected collocates for the node, where distractors were high	Semantically non-transparent Verb + Noun collocations with low L1 congruency side-by-side with their	Four groups of EFL university-level students whose L1 was Taiwanese,

Studies	Methodologies	Test Items	Participants
(2013)	frequency words from the General Service List (West, 1953); 2. Treatment in which participants read and listened to a story with different number of repetitions of target collocations (1, 5, 10, and 15 times); 3. Four post-tests where participants: 1. Provided a collocate for the node; 2. Completed a multiple-choice test similar to the pre-test; 3. Translated from L1 to L2; and 4. Translated from L2 to L1	L1 translations, selected on the basis of their t-scores (strength of association) and high frequency words from the General Service List (West, 1953)	with a minimum of 6 years of formal instruction; proficiency level was measured by the Vocabulary Levels Test (Schmitt et al., 2001);
Edmonds and Gudmestad (2014)	1. Collocation judgement + multiple-choice test where participants identified adjectives that could match an adverb; 2. Fill-in-the-blank test where participants completed sentences by providing an adverb for the adjective; 3. Questionnaire on participants' demographic information and language	Adverb + Adjective collocations selected on the basis of their frequency and t-scores	NNSs (university students majoring in English whose L1 was French) and NSs

Studies	Methodologies	Test Items	Participants
	background		
Peters (2016)	1. Pre-test where participants translated L1 collocations into L2, where the first letter of a collocation in English was provided; 2. Fill-in-the-gap and matching exercises involving explicit form-focused instruction, where the L1 translation of collocations was provided; 3. Two form recall tests where participants translated L1 collocations into L2 with and without the node word in English already provided; 4. Form recognition test where participants matched nodes and collocates	Verb + Noun and Adjective + Noun semi-restricted congruent and incongruent collocations with the nodes retrieved from the high frequency lists in COBUILD English Dictionary for Advanced Learners (Sinclair, 2001), with the length ranging between 9 and 17 letters	ESL speakers whose L1 was Dutch, with 4-6 years of prior instruction
Szudarski and Carter (2016)	1. Treatment: Participants read stories with collocations underlined/not, where the number of encounters was 6 to 12;	Verb + Noun collocations consisting of frequent verbs such as *make, take*, and *have* and infrequent collocates;	Two groups of EFL learners whose L1 was Polish, with a minimum of 6 years of

Studies	Methodologies	Test Items	Participants
	2. Recall/recognition/ production-based post-tests where participants: (a) Translated Verb + Noun collocations from L1 to L2; (b) Provided an adjective for a noun node, where the first letter and the number of letters in the word was provided; (c) Provided a verb for a noun collocate, where no first letter or the number of letters in the word was provided; (d) Translated Adjective + Noun collocations from L2 into L1; and (e) completed a multiple-choice test where participants selected a verb for the Verb + Noun collocations	and Noun + Adjective collocations with frequent adjectives and infrequent nouns retrieved from the BNC; selected on the basis of their high MI score	formal instruction; proficiency level was measured by the Vocabulary Levels Test (Schmitt et al., 2001)
Zhang (2017)	1. Receptive collocational knowledge pre-test where participants translated collocations from L2 to L1; 2. Treatment: (a) Reception-focused, where participants read collocations that were highlighted in sentences	Verb + Noun collocations retrieved from Ackermann and Chen's (2013) Academic Collocation List based on their length, degree of semantic transparency and congruency, and alliterations; sentences retrieved	Four groups of ESL learners whose L1 was Chinese, divided according to the types of treatment they have been receiving; proficiency level was measured by the

Studies	Methodologies	Test Items	Participants
	alongside L1 translations; (b) Production-focused, where participants wrote sentences with collocations based on the L1 translation; (c) Reception + production-focused, where participants read collocations in sentences along with L1 translations and wrote sentences in English using collocations; 3. Six post-tests and delayed post-tests where participants (a) Provided collocates for a node; (b) Completed a multiple-choice test, where they chose the correct collocate for a node; (c) Translated from L1 to L2; (d) Translated from L2 to L1; (e) wrote sentences using collocations; (f) Completed a multiple-choice test, where they identified a sentence in which to use the correct grammatical form of a collocation	from the BNC and Oxford Advanced Learner's English-Chinese Dictionary (2004)	Vocabulary Levels Test (Schmitt et al., 2001)

In spite of the variety of factors that are examined, it is possible to notice that most studies mostly use standardized formats that are usually implemented when testing single words (Carter & McCarthy, 1988; Schmitt et al., 2001), for example, multiple-choice, matching, or fill-in-the-blank. Even those researchers who focus on designing innovative test formats (Eyckmans, 2009; Jaén, 2007) still rely on these kinds of traditional structures. This is a reasonable solution since standardized formats are less time-consuming and facilitate test validation, i.e., establishing construct validity and reliability. However, a significant drawback is that such tests are targeted at individual collocation components rather than whole collocations. Specifically, most of the tasks ask participants to provide/select either a node or a collocate, which contradicts the principle of holistic processing and usage of collocations and does not take into account strength of association and semantic/structural restrictions. Translation-focused tasks are a slightly better fit in terms of addressing whole collocations; however, they are mostly measuring productive knowledge and over-rely on L1-L2 interaction that will trigger collocational errors related to the factors of L1 interference and congruency. As a result, it is not an ideal condition for evaluating collocational competency in L2.

In the abundance of research on testing/teaching collocations, one study stands out as addressing whole collocations and not relying on L1: Revier (2009). Revier's study provides a new format of recognition/controlled production-based test, which, although multiple-choice-based, asks participants to choose both the node and collocate from the grid options and pay attention to the context of the sentence with the gap. However, all of the distractors can potentially be combined into English-like collocations, and so the context remains the only factor that guides participants in their choice. This circumstance

does not provide any insights into internal collocation-specific factors that might trigger recognition and controlled production.

Another significant limitation of the studies listed in the tables is the focus on one or two morphosyntactic collocation types only: Adjective + Noun or Verb + Noun collocations. Researchers justify their preference for these particular types of collocations by their frequency, usefulness, and convenience of retrieving them from the corpora. However, such an approach does not examine the possible connection that might exist between the morphosyntactic structure and collocation acquisition. Gitsaki (1996) is one of the few researchers who moved beyond the standard approach when combining several morphosyntactic collocation types after Benson et al. (1986). Although the approach Gitsaki used to select morphosyntactic structures is somewhat inconsistent since collocations are mixed with other formulaic sequences (for example, phrasal verbs), she identified some statistically significant differences; for example, lexical collocations are easier to recognize and produce than grammatical collocations.

Summary

This section of the literature review examined the research methodologies designed to test collocation recognition, recall, and controlled production. While the research objectives, design, characteristics of the testing items, and target populations varied, most studies used standardized multiple-choice, translation, or matching tasks, which either do not adequately measure knowledge of whole collocations or over-rely on L1. Two test formats appear to be appropriate in terms of addressing whole collocations and tapping into such aspects of collocational competence as form, meaning, and use (Webb et al., 2013), as well as intuition and fluency (Wray, 2008): one is the

acceptability judgment task format and the other is the CONTRIX grid-based test

(Revier, 2009). Both allow participants to choose a whole correct collocation rather than

its individual components. However, while most studies that use acceptability judgement

tasks focus on context-free word combinations, CONTRIX presents sentences with a

cloze and, thus, test items are tied to a specific context. In terms of the distractor

selection, it is important to address such internal collocation-specific factors as the

frequency of occurrence and frequency or co-occurrence (MI score) of the items. The

principle of synonyms used by Koya (2005), Nizonkiza (2015), and Webb and Kagimoto

(2011) is also promising since it allows tying collocations to a specific context.

Chapter 5: Stage I. External Factors

Rationale and Design

Different theoretical and empirical studies have shed light on the external factors influencing collocation acquisition and the difficulties that L2 speakers encounter when recognizing and producing formulaic sequences. Several researchers have explored extralinguistic factors such as the age of onset of learning English (Granena & Long, 2013; Wray, 2002), the absence of native speaker intuition (Gitsaki, 1996; Siyanova, 2010; Siyanova-Chanturia & Martinez, 2014; Siyanova & Schmitt, 2008), the length of residence in an English-speaking country (Foster, 2009; Groom, 2009; Schmitt, 2000), personality and learning habits (Adolphs & Durow, 2004; de Wit, 2007; Wray, 2008), the frequency and quality of language input and output (Cieślicka, 2015; Szudarski & Carter, 2016), vocabulary learning strategies (Lewis, 2000; Mian, 1988; Wray, 2002), and reading strategies (Sonbul, 2015). Other researchers have focused on interlinguistic factors, such as the predominant language for thinking and communication (Wang & Shih, 2011; Xu, 2015), L1-L2 congruency (Bylund, Abrahamsson, & Hyltenstam, 2012; Liao, 2010; Millar, 2011), and level of language proficiency (Groom, 2009; Li & Schmitt, 2010; Nizonkiza, 2012, 2015; Talakoob & Koosha, 2017). However, only a few studies by Bardovi-Harlig and Bastos (2011), Carrol et al. (2016), Schmitt et al. (2004), and Yamashita and Jiang (2010) have attempted to examine and compare the significance and relevance of multiple factors in comparison to one another in order to create a comprehensive typology. Consequently, Stage I of this research was designed to address this problem by identifying the patterns and classifying/comparing different extralinguistic and interlinguistic factors in order to answer the following research

question: What external factors impact recognition of false collocations by speakers of English as L1 and L2?

This study focused on participants' ability to recognize collocations since recognition as a process of selecting an appropriate and correct word combination from the set of possible options was identified as the basis for collocational competence and, therefore, subsequent collocation production (Gyllstad, 2009; Henriksen, 2013). Additionally, recognition can be measured in a more reliable way through standardized tests (Eyckmans, 2009; Gyllstad, 2009). The decision to measure and analyze participants' recognition of false (incorrect and non-English-like) collocations was motivated by the fact that, according to Gitsaki (1996), Schmitt (2010), and Wray (2002), the L2 learners' predominant collocational errors are related to accuracy of collocational use, which includes paraphrasing, blending, or interchanging collocational components. This means that unlike speakers of English as L1, ESL learners cannot effectively recognize the holistic nature of formulaic sequences (Cieślicka, 2015; Henriksen & Stoehr, 2009; Wray, 2002) and consequently produce a large number of false collocations. Therefore, the assumption of Stage I of the research is that by testing participants' ability to recognize these incorrect collocations and correlating test scores with a number of independent variables, those external factors that might be more or less influential could be identified.

In order to identify and compare these factors, first a pilot study was conducted that dealt with the question of what criteria led participants to recognize false collocations. A recognition-targeted test which combines two formats –an acceptability judgement task popularized by Siyanova and Schmitt (2008), Wolter and Gyllstad (2011;

84

2013), Wolter and Yamashita (2015), and Yamashita and Jiang (2010), and a multiple-choice exercise used by Gyllstad (2009), Nizonkiza (2015), and Ngyuen and Webb (2016)– was developed. A post-test interview was then conducted to identify patterns that could point to potentially significant factors. Second, a large-scale quantitative study that included the first part of the same test (an acceptability judgement task) and a post-test survey delved deeper into those factors that had been identified as significant in the pilot study.

Both studies focused on the differences and similarities of false collocation recognition by the two groups (speakers of English as L1 and L2) and what factors might be potentially responsible for these differences and similarities. The rationale behind employing both L1 and L2 speakers was the necessity to compare the two perspectives in order to understand the mechanisms of the ESL learners' collocation acquisition. Thus, Stage I partially replicated the angle of research by Gyllstad (2007, 2009), Siyanova and Schmitt (2008), Wolter and Gyllstad (2011, 2013), Wolter and Yamashita (2015), and Yamashita and Jiang (2010), who explored specific mitigating and/or predicting factors influencing collocation recognition by comparing NSs and NNSs of English.

Pilot Study

Participants (see Table 5). The participants in the pilot study were five college and university students from different genders, L1 backgrounds, and aged from 21 to 26. One was a speaker of English as L1 and the other four were speakers of English as L2. Since it was impossible to adequately measure the ESL speakers' proficiency level without performing a series of proficiency tests, the study implemented two procedures to differentiate these four participants. First, they were divided into emergent and advanced

bilinguals based on the age of onset of learning English and how long they had been

living and studying in an English-speaking country. Emergent bilinguals are ESL

learners, in particular immigrants or international students, who started learning English

somewhat late, have spent less than five years in an English-speaking country (Foster et

al., 2014; Lahmann et al., 2016), and have not yet achieved full bilingual competence

because their L1 still dominates in their oral and written communication (Menken, 2013).

Advanced bilinguals are learners who started learning English earlier and have continued

their studies over a longer period of time. As a rule, they have spent more than five years

in an English-speaking country and their English proficiency level might become

dominant and/or approach their L1 proficiency level (Office for Standards in Education,

Children's Services and Skills [Ofsted], 2005; Department for Education and Skills

[DfES], 2006). Second, self-assessment was implemented. Participants were asked to

self-evaluate their proficiency based on a Likert scale from 1 ("poor") to 5 ("excellent").

Table 5

Pilot Study Participants

Participants	English language proficiency	Age of onset of learning English	Formal instruction	Length of residence in an English-speaking environment
Interviewee X	Monolingual speaker of English as L1 (at some point attempted to learn French as L2)	Not applicable	Not applicable	Not applicable
Interviewee Y	Advanced bilingual (Russian/English); Self-assessment on the Likert scale was 5	7	Mandatory English classes at school	One year of studying in the USA; immigrated to Canada at the age of 17; by the time of the interview, had lived in Canada for 10 years
Interviewee Z	Emergent bilingual (German/English); Self-assessment on the Likert scale was 2-3	14	Mandatory English classes at school	Came to Canada for 6 months as an exchange student; by the time of the interview, had lived in Canada for 4 months
Interviewee U	Emergent bilingual (German/English); Self-assessment on the Likert scale was 3-4	10	Mandatory English classes at school	Did an internship in England for half a year; came to Canada for a year as an exchange student; by the time of the interview, had lived in Canada for 5 months
Interviewee W	Emergent bilingual (Chinese/English); Self-assessment on the Likert scale was 3-4	12	Mandatory English classes at school	Came to Canada at the age of 18 as a student; by the time of the interview, had lived in Canada for 2.5 years

Instruments. The recognition-based test consisted of two parts. The first part was an acceptability judgement task (see Appendix C) partially modelled after Siyanova and Schmitt (2008), Szudarski and Conklin (2014), and Yamashita and Jiang (2010). This task included five neutral register paragraphs adapted from Canadian newspapers and magazines such as the Canadian Press, the Globe & Mail, and the CBC. These abridged and simplified paragraphs addressed everyday topics (e.g., weather forecasts and news reports), and the amount of specialized vocabulary was minimal. The whole text was approximately 500 words long. The test items (see Appendix D) were 18 Adjective + Noun and Verb + Noun general English collocations that consisted of lexical bigrams, that is, two consecutive words (Granger & Bestgen, 2014; Vechtomova et al., 2003), in some cases connected by a functional preposition and/or an article. These morphosyntactic types are the most frequently addressed in existing research (Boers et al., 2014; Eyckmans, 2009; Gyllstad, 2009; Jaén, 2007; Koya, 2005; Pellicer-Sánchez, 2017) and are believed to cause more challenges for ESL speakers. In each of the 18 test items, one of the two lexical words was deliberately changed to its synonymous distractor because, as Wray (2002) and Cieślicka (2015) suggest, ESL learners often paraphrase collocations, that is, substitute words by their synonyms, which are inacceptable in a given context. The synonymous distractors were run through the BNC in order to make sure that they do not co-occur with their respective nodes.

The second part of the test was a traditional cloze and multiple-choice test (see Appendix E) similar to the tests used by Gyllstad (2009), Ngyuen and Webb (2016), and Nizonkiza (2015). It included a set of 10 sentences retrieved from the TIME Magazine Corpus. Each sentence had a blank (a missing part of a collocation). Three choices (one

correct word and two synonymous distractors) were provided below the sentence. As in the case of the acceptability judgement task, test items were Adjective + Noun and Verb + Noun collocations and consisted of lexical bigrams.

The rationale behind using these two different test models was the necessity to identify an optimal model that would effectively measure recognition of false collocations. The two models were selected because, according to the findings of Chapter 4, both are frequently used in studies on collocation recognition. Three speakers of English as L1 piloted the test and provided their judgement on the acceptability of the test items and distractors and the difficulty and comprehensibility of the text.

The semi-structured post-test interview questions (see Appendix F for sample questions) were divided into four main parts. The first part addressed participants' immediate post-test experience:

Reading strategies. A set of questions asked how they identified "odd-sounding" or "non-English-like" patterns, and what they focused on (word form or meaning, adjacent words, position and function in a sentence, etc.). These questions helped to identify the strategies that participants used to complete the task and recognize false collocations.

Test items. Another set of questions asked what word combinations seemed challenging/easy, and what word choices participants were not sure about. These questions provided information about which test items were effective in measuring collocation recognition and those that were problematic and had to be eliminated from the subsequent quantitative study.

Test environment. A final set of questions asked if the time for completing the tasks was sufficient, and which of the two tasks was more/less difficult. These questions helped to evaluate and compare the test tasks in terms of their difficulty and effectiveness and helped to set time limits for the subsequent quantitative study.

The second part of the interview examined participants' demographics and language background (for speakers of English as L2, this included at what age they started learning English, for how many years, what additional languages they knew). The third part of the interview was for speakers of English as L2 only and focused on participants' overall English language proficiency in terms of their self-assessment of the four language skills (reading, listening, speaking, and writing). The fourth part of the interview addressed their vocabulary-learning strategies focusing on how they used/preferred to learn word combinations.

Data collection procedures. Participants read the text one time and underlined 2-to-3-words-long "non-English-like" combinations. In order to ensure spontaneity of recognition, no explicit explanation of collocations was provided. Based on Ziefle (1998), the average reading speed is 200 words per minute. However, since participants were focusing on specific details, they were given slightly more time: 20 minutes (approximately 1.5-2 minutes per one paragraph in Task 1 and one minute per one sentence in Task 2). Although emergent bilinguals read more slowly, all five participants finished the test in 20 minutes and indicated that they read it either "at their normal speed" or "slower". The semi-structured interview was conducted immediately after the test.

Data analysis and results. Whereas for the multiple-choice test the total number of correct responses was calculated, for the acceptability judgement task three response categories were identified: (1) false collocations correctly identified by the participants; (2) word combinations (including missed false collocations) that participants wanted to underline but did not; and (3) word combinations that were actually correct but considered wrong by the participants. When the test takers were asked to explain their choices, they mentioned intuition, prior knowledge of a word combination, and attempts to analyze links between words.

After calculating the percentage of correct responses, data analysis focused on the five interview transcripts (see a sample transcript in Appendix F). While coding the interviews, structural and descriptive coding procedures were implemented (Saldaña, 2012). First, structural coding built upon the four categories of interview questions, namely, test environment and test-taking strategies, participants' demographics, prior language learning, and self-assessment of their English language proficiency. This step allowed segmenting the interview data into several general sections and identifying general patterns. Second, descriptive coding focused on the patterns of extralinguistic and interlinguistic factors potentially influencing collocation recognition that had already emerged as significant in the existing research as discussed above in the section "Rationale and Design".

Test scores on the first part of the test (acceptability judgement task) were higher for the monolingual speaker of English (83% for Interviewee X) and the advanced bilingual speaker of Russian/English (72% for Interviewee Y) and lower for the emergent bilinguals, which is evidence that language proficiency and English as L1 impact

collocation recognition. The scores varied among the emergent bilinguals, where the difference was not visibly impacted by the length of their residence in an English-speaking country; for example, Interviewee Z had the shortest stay in Canada (4 months) and the highest score (55.5%), while Interviewee W had the longest stay (2.5 years) and the lowest score (22%). The scores also did not seem to depend on the age of onset of learning English: for example, Interviewee Z, who had the highest score, had started learning English later than Interviewees W and U. Self-assessment of English language proficiency did not seem to impact test scores, either: for example, Interviewee Z gave herself a score of 2-3 out of 5; while Interviewees U and W gave themselves higher scores.

The scores on the second part of the test (cloze and multiple-choice test) were also higher for the advanced bilingual and the speaker of English as L1 (100%) than for the emergent bilinguals (70%). Additionally, the scores were generally higher than those of Task 1. Therefore, the acceptability judgement task seems to be more difficult than the multiple-choice exercise, which corresponds to findings from Wolter and Gyllstad (2011, 2013), Wolter and Yamashita (2015), Yamashita and Jiang (2010), and other researchers who suggest that NNSs cannot provide accurate judgements on whether word combinations are used in English or not.

An analysis of emerging patterns and their correlation to test scores identified several groups of factors that might have influenced participants' recognition of false collocations: (1) extralinguistic factors (language input; cognitive factors, such as vocabulary-learning strategies, focus of attention while reading (criteria of "word choice"), and attention span); and psychological-affective factors of the communicative

situation, such as necessity, motivation, and attitudes towards language learning); (2) interlinguistic factors (L1-L2 language interinfluence; code-switching; and predominant language for communication); and (3) test task factors. These categories are described in more detail below.

Extralinguistic factors.

Frequency and quality of the input. The four speakers of English as L2 confirmed that they became more fluent in English after coming to Canada as a result of interacting with speakers of English as L1. Those participants who were taking ESL classes stated that classroom instruction was different in Canada and their home countries.

> Interviewee Z: *At the airport I [couldn't] ask stewardess for a drink because I didn't know how to say it. It is a completely other situation. Because I just spoke English in my classroom or I watched movies in it, but I never had a real conversation[1]. <...>.*

Vocabulary learning strategies. The emergent bilinguals with higher test scores mentioned that previously, their formal language instruction focused on learning words one-by-one in thematic lists; however, upon coming to Canada they realized the importance of "expressions" and context.

The focus of attention while reading (criteria of "word choice"). While completing the acceptability judgement task, participants implemented different criteria to identify non-English-like word combinations. For example, while the speaker of

---

[1] In this and subsequent quotes, participants' original sentence structure, wording, and grammatical patterns were preserved.

English as L1 and the advanced bilingual focused on the "structure of the word combination" and overall "sentence flow", emergent bilinguals focused on the meaning of individual words.

Attention span. Participants reported different preferences for reading the text. For example, those participants who had higher test scores read for "the fluency of the whole sentence", while participants with lower test scores read word-by-word.

Communicative situation. Participants reflected on different situations when they felt more or less comfortable communicating in English, where usually more stressful and formal situations caused them to interact less and make more mistakes.

Interviewee W: [In class] *only you not English speaker. It's kind of awkward and so nervous if you stand in front of all speaking English people, speaking English in an academic way.*

Interviewee Z: *I was very shy and I didn't talk a lot. I was afraid of making mistakes and the others would judge about my English and my accent.*

Need, motivation, and attitudes towards language learning. While four ESL learners were highly motivated to learn English and going to study and/or live in Canada gave them an additional stimulus of the necessity of learning, the speaker of English as L1 did not experience a strong need to learn an additional language (French). As a result, the four ESL learners pushed themselves to communicate in English exclusively upon their arrival to Canada, which accelerated their learning progress.

Interviewee Z: *I have to communicate if I want to express myself, and if I want to tell something that I have no option. I am forced to speak English.*

On the contrary, the speaker of English as L1 admitted that he did not need French in his daily life, indicating a low level of motivation and necessity. At a certain point, he noticed that he was lagging behind in his studies and so, feeling that he would not be able to achieve the same proficiency level as NSs of French and his more advanced classmates, he dropped out.

Interlinguistic factors.

L1-L2 interinfluence and code-switching. The four ESL speakers admitted that they often relied on their L1 when speaking or writing in English; however, the extent of L1 influence differed based on their proficiency in English. For the emergent bilinguals, their L1 interfered with English on a regular basis, whereas the advanced bilingual reported that language interference for him was two-fold: while his mother tongue, Russian, influences his writing in English ("positioning of certain words or sentences"), English often influences his speaking in Russian:

> *...when I talk I would start a sentence in Russian, but the ending of the sentence would be English. <...> When I am listening to myself and I am figuring out that it doesn't sound right, I just switch to English.*

There was a positive relationship between L1-L2 code-switching and the length of residence in an English-speaking country. Interviewee Z (emergent bilingual) almost never code-switched ("It really hurts my brain"), while other ESL speakers code-switched more often while speaking and sometimes writing (informal emails).

Language of communication. For the emergent bilinguals, their L1 remained their primary language for interaction. However, their daily usage of English varied based on their language proficiency, length of residence in an English-speaking country, and the

95

quantity and quality of English input and output, particularly communication with speakers of English as L1. For example, Interviewee Z, who had spent 4 months in Canada and predominantly socialized with her German speaking friends, said that she used to mentally translate from German to English. Interviewee U, who spent 6 months in England and preferred to socialize with speakers of English as L1, stated that she communicated in English 50% of the time. Meanwhile, Interviewee W, who had spent in Canada 2.5 years and had both English and Chinese-speaking friends, communicated predominantly in English (70% of the time). The advanced bilingual pointed out that every year his use of Russian progressively decreased, where by the time of the interview, he used English even while communicating with his Russian-speaking friends.

The factor of the test task. All five participants indicated that the cloze and multiple-choice test was easier because of the limited options of choice, which corresponded to the higher test scores on this task. However, their opinions differed on the acceptability judgement task. While the speaker of English as L1 and the advanced bilingual stated that the task was not particularly difficult and once they started reading the text, non-English-like word combinations were obvious, emergent bilinguals admitted that they were confused and not sure what to look for in the text. This confirms the findings of Siyanova and Schmitt (2008), Wolter and Gyllstad (2011, 2013), Wolter and Yamashita (2015), and Yamashita and Jiang's (2010) studies on ESL learners being unable to differentiate between acceptable and non-English-like collocations.

Interviewee U (emergent bilingual): *The test was more difficult than the... selecting the right word for the sentences cause I wasn't sure.*

The final part of the analysis focused on the test items. Participants were asked to comment on their choice of non-English-like word combinations to help identify the criteria that guided them in their recognition of false collocations. The criteria are outlined in Table 6.

Table 6

Criteria for Recognizing False Collocations

Criteria	Participants	Comments
Context	All five	Interviewee Z: *I know what it means but in this context for me it doesn't make sense.* Interviewee Z: *It seem to be wrong word in this context. Maybe, 'unlucky' or another word in general [about the word combination 'luckless incident'].*
Background knowledge	All five	Interviewee W: *I don't usually say like... in that way.* Interviewee X: *I hear that on the news 24 hours a day.<...> I never see it used.*
Formal/ informal style of a text	Speaker of English as L1 and two more advanced speakers of English as L2	Interviewee Z: *I don't know... is it just a slang, or can you use that in this context? <...> Yeah, is it very formal, or is it informal?* Interviewee Y: *'Fluctuated in age' is too scientific. It just talks about children at school... Must be something different. Maybe, different age.*
Intuition	Speaker of English as L1 and advanced bilingual	Interviewee X: *I chose them I guess because they didn't fit to... didn't combine. One can make a better choice... It all sounds as if spoken by a non-*

Criteria	Participants	Comments
		native speaker.
		Interviewee Y: *Just feeling tells you there was something wrong, but I could not 100 percent say what exactly.<...> I rely more on my feeling probably, not my knowledge.*
Analysis	Advanced bilingual	Interviewee Y: *The positioning in the sentence; the structure of the word itself. Like this[Good-being]... It is not wrong but obviously the saying is well-being.*
		Interviewee Y: *You don't say 'unimportant injury'. It's heavy, or bad, or small.*
Language comparison/ Translation	Two speakers of English as L2	Interviewee Y: *I just read it in my head, and some of the word combinations I know are supposed to sound in a different way. If you probably translate them directly into Russian in my case, they would make sense, but it's not in English, the word combination.*
		Interviewee Z: *I tried to translate it. <...> Yeah, in my mind, I translated it to my mother tongue.*

Although most of these criteria (for example, relying on context and using background knowledge) were common for all interviewees, the use of certain strategies (for example, using intuition, analyzing words, or comparing languages) varied. For example, only the speaker of English as L1 and advanced bilingual relied on their intuition. Although emergent bilinguals might have felt something was wrong with a particular word combination, they usually did not trust their "gut" feeling. This uncertainty becomes obvious in their responses.

Interviewee U: *As I am a not native speaker, it is difficult to yeah... say whether it's just my English, my English knowledge, or is the sentence could be like this?*

Interviewee Z: *Sometimes I was a little bit unsure but just if this sounds just awkward to me because I don't know that word or because I was right.*

Instead, ESL speakers implemented analysis and language comparison. This finding corresponds to the majority of studies by Nation and Newton (1997), Siyanova (2010), Siyanova-Chanturia and Martinez (2014), Schmitt (2010), and Wray (2002), who suggest that ESL speakers tend to analyze word combinations rather than take a guess.

Discussion. This pilot study of five participants (one speaker of English as L1 and four advanced and emergent bilingual speakers of English as L2) started with a quantitative analysis of test results. The test was targeted at recognition of false collocations and used acceptability judgement and cloze multiple-choice task formats. Then, a qualitative structural and descriptive coding-based analysis of semi-structured post-test interviews with participants was conducted. The results of these analyses provided a broad picture of factors that had been previously identified in other scholarly work (as summarized in Chapter 3) as potentially influential on collocation recognition. These factors include:

1. Extralinguistic factors: frequency and quality of language input; cognitive factors such as vocabulary learning strategies, the focus of attention while reading (criteria of "word choice"), and attention span; and psychological-affective factors of the communicative situation, such as necessity, motivation, and attitudes towards language learning;

2. Interlinguistic factors: L1-L2 language interinfluence; code-switching; and the predominant language of communication;

3. The factor of the test task.

At the same time, some factors that were significant in previous research by Granena and Long (2013), Schmitt (2000), and Wray (2002), such as the length of residence in an English-speaking country or the age of onset of learning English, did not appear to influence test scores. This contradictory finding indicates that although some trends and patterns emerged, it is not possible to determine what influences collocation recognition based on test scores and interviews with only five participants. More data are needed to arrive at generalizable conclusions.

The potentially significant factors manifested differently depending on whether participants were speakers of English as L1 or L2 and/or advanced or emergent bilinguals. This finding corresponds to studies by Ellis et al. (2008), Sonbul (2015), and Wolter and Gyllstad (2013) that similarly argue for language proficiency as the primary overarching factor of influence. For example, this study identified that cognitive factors, such as the focus of attention while reading (criteria of "word choice") and attention span while reading, differ: while the advanced speaker of English as L2 and the speaker of English as L1 focused on the overall sentence structure and flow, emergent bilinguals focused on the meanings of individual words. The reading span was also different: while the advanced speaker of English as L2 and the speaker of English as L1 attempted to read whole sentences, the emergent bilinguals with lower test scores focused on individual words. A possible explanation can be found in studies by Ellis et al. (2008), Hill, (2000), Schmitt (2013), Siyanova and Schmitt (2008), and Wray (2002) that suggest that L2

speakers process formulaic sequences more slowly and therefore have a lower reading fluency.

According to the analysis of test results and interview patterns, it can be said that the advanced speaker of English as L2 and the monolingual speaker of English as L1 scored higher on both test tasks because of their higher English language proficiency level, which corresponds to the findings of Laufer and Waldman (2011), Schmitt (2010), and Wray (2002) who stated that collocational competence increases as language proficiency improves. It is important to note that those speakers of English as L2 who had higher test scores had formed specific communicative and cognitive habits. For instance, they mostly used English for communication (Wang & Shih, 2011), code-switched more often (Cieślicka, 2015), learned new words in combinations rather than one-by-one (Gitsaki, 1996; Lewis, 2000; Wray, 2002), and were motivated and willing to engage in interaction with speakers of English as L1 (de Wit, 2007).

The criteria that guided participants in selecting non-English-like collocations in the acceptability judgement task were also different for the speaker of English as L1, the advanced speaker of English as L2, and emergent bilinguals. Although all participants paid attention to context and tried to activate their background knowledge, as confirmed in studies by Cieślicka, (2015) and Webb et al. (2013), only the advanced speaker of English as L2 and the speaker of English as L1 relied on their intuition when providing acceptability judgements on the word combinations. This finding corresponds to the results of previous studies (Siyanova & Schmitt, 2008; Siyanova-Chanturia & Spina, 2015; Wray, 2002, 2008) that indicated that ESL learners do not have a well-developed

intuition and are not always sensitive to the semantic and morphosyntactic links between collocational components.

Finally, the significant difference in the test scores on the two tasks (the average score on Task 1 was 58, and the average score on Task 2 was 82) suggests that cloze and multiple-choice tests cannot effectively measure collocation recognition because they give limited choice options (two of which are obviously false), which automatically focuses participants' attention on the correct choice and does not measure their knowledge of the whole collocation (Revier, 2009). Even those emergent bilinguals who scored quite low on the acceptability judgement task had comparatively higher scores on the cloze and multiple-choice test and explained during interviews that the latter task was easier because it gave them more options for taking a guess.

On the whole, the results of the pilot study identified some external factors that have already been explored in other research studies, as discussed in Chapter 3. However, it remained to be seen whether these factors are statistically significant. For this purpose, a large-scale quantitative study followed the pilot study to further explore the external factors impacting recognition of false collocations by speakers of English as L1 and L2.

Quantitative Study

This large-scale quantitative study had two objectives. First, it analyzed and compared those external factors impacting recognition of false collocations that had been suggested as potentially significant in the pilot study in order to create a comprehensive typology. Second, since the pilot study identified that the factors of English as L1 and English proficiency level might be overarching in the framework of external factors, this subsequent quantitative study compared the two populations of participants (speakers of

English as L1 and L2) in order to examine how other external factors correlated with participants' English language proficiency, and what the differences and similarities between groups might be.

Participants. As displayed in Table 7, there were a total of 93 participants in this large-scale study: 50 speakers of English as L1 and 43 speakers of English as L2. The majority of participants were in the age range of 18-19 and almost two thirds were male. In the group of speakers of English as L2, most participants spoke Arabic (n = 11 [25.5%]) or Chinese (n = 17 [39.5%]) as their L1 and had started learning English between the ages of 5 and 12 years old (n = 27 [62%]). On average, participants had spent less than five years in an English-speaking country (n = 34 [79%]).

Table 7

Participants: Age, Gender, and Languages

	English as L1	English as L2
Total number of participants	50	43
Age range	17-29	16-38
Predominant age group	18-19 (n = 40)	18-19 (n = 22)
Gender	36 males and 14 females	30 males and 13 females
Additional/primary languages	31 participants spoke French; 2 participants spoke Arabic; and one participant spoke Chinese	17 participants spoke Chinese; 11 participants spoke Arabic; 4 participants spoke French; and 11 more participants spoke other primary languages (e.g., German,

	English as L1	English as L2
		Hindi, Portuguese, and Urdu)
Age of onset of learning English	Not applicable	1 – 32 (the majority started learning English between 5 and 12 years old; n = 27)
Length of residence in an English-speaking country	Not applicable	4.35

Instruments. Participants were administered the first part of the test (acceptability judgement task, see Appendix C) that was used in the pilot study. The rationale for focusing on the acceptability judgement task was the finding of the pilot study that cloze and multiple choice tests did not measure the whole collocation recognition adequately, as it provided only limited choice options and did not allow for the spontaneity of recognition. To reiterate the pilot study test, the short paragraphs retrieved from Canadian newspapers with deliberately inserted 18 incorrect collocations represented samples of neutral general English describing different aspects of everyday life. No knowledge of specific terminology was required. The 18 tests items belonged to the two most widely used morphosyntactic types: Adjective + Noun and Verb + Noun (Boers et al., 2014; Jaén, 2007; Koya, 2005; Pellicer-Sáncez, 2017) and presented short-span consecutive two-word lexical combinations (Vechtomova et al., 2003) to ensure immediate recognition.

A post-test survey (see Appendix G) was administered immediately after the test. When designing the survey questions, the focus was on both the findings that had already emerged from the pilot study and the existing studies of factors influencing formulaic

language acquisition as discussed in Chapter 3. The first part of the survey focused on participants' demographics, particularly their language background. The second part targeted their post-test experience and the strategies they employed both for reading the text and selecting non-English-like word combinations. The third part addressed participants' predominant language of communication and L1-L2 interinfluence. The survey consisted of closed-ended questions (multiple choice, Likert scale, and yes/no questions); however, participants were encouraged to provide additional comments below each question.

Although the question patterns were similar, the questions themselves were not identical for speakers of English as L1 and L2. While questions for the ESL speakers collected information on the number of years they have spent learning English (in an English-speaking country and their home countries), their self-assessment of their English language proficiency, vocabulary-learning strategies, and their predominant language of communication, questions for speakers of English as L1 focused on any additional languages they might know in order to examine potential L1-L2 interference.

Data collection procedures. Participants were instructed to read the text once only and then underline 2-3-words-long non-English-like ("odd-sounding") phrases. They had approximately 8 minutes (1.5 minutes per one paragraph) to complete the task, as this timing had been sufficient in the pilot study for the speaker of English as L1 and the advanced and emergent bilinguals alike. As in the pilot study, participants were not given any explanation on collocations in order to ensure spontaneous recognition of false collocations. Immediately after taking the test, participants had 15 more minutes to complete the survey.

Data analysis and results. The data were coded using SPSS Version 22.0. Unlike

the pilot study that examined three response categories (false collocations that were

correctly identified by the respondents, missed false collocations, and actually correct

word combinations that were underlined by the participants as wrong), this large-scale

quantitative study focused only on the number of correct responses, that is, how many

false collocations participants identified out of a possible 18. The frequency counts and

descriptive statistics for the test scores and each of the independent variables showed that

kurtosis and skewness were within the norm (+/- 2), which indicated the normality of the

data, as per Trochim and Donnelly (2006), Field (2009), and Gravetter and Wallnau

(2014). Several analytical procedures were then implemented to compare the total test

scores of speakers of English as L1 versus those of speakers of English as L2, and

explore relations between the dependent variable (test scores) and several independent

variables that might have been similar or different across the two groups of English

speakers.

　　　Step 1. Relationship between total test scores and English as L1 and L2. An

independent samples t-test was conducted to compare the means of a normally distributed

continuous dependent variable (test scores) across the two groups (speakers of English as

L1 and as L2). The difference was statistically significant in the scores for speakers of

English as L1 (*M* = 49.8, *SD* = 21) and L2 (*M* = 27.5, *SD* = 16.7), where *t* (93) = 5.67 and

p = .000. The Sig. value (.18) was above the cutoff of .05, which means that the variances

for two groups were assumed as equal. The magnitude of the differences in the means

between the two groups was large (eta- squared statistics = .24) This finding suggests that

being a speaker of English as L1 has a significant impact on recognition of false

collocations, which explains why speakers of English as L1 had higher scores on the collocation recognition test in comparison to speakers of English as L2. The finding is in agreement with studies by Groom (2009), Li and Schmitt (2010), Siyanova and Schmitt (2008), Wolter and Gyllstad (2011, 2013), Wolter and Yamashita (2015), Wray (2002), and Yamashita and Jiang (2010), who suggested that speakers of English as L2 recognize and use collocations less fluently and accurately than speakers of English as L1.

Step 2. The factor of English as L1 and test-taking strategies. A chi-square test of association was run for both groups of participants to explore if the factor of English as L1 or L2 might have a statistically significant relation to the categorical variables of the test-taking strategies that the participants employed, such as reading speed, the focus of attention while reading (criteria of "word choice"), and attention span.

The results showed that there was no statistically significant association between the strategies that participants implemented to read the text and the factor of English as L1 or L2: for reading speed [$x(1) = 0.78$, $p = .670$], for focus of attention [$x(1) = 7.06$, $p = .070$], and for attention span [$x(1) = 0.15$, $p = .980$]. This means that the majority of participants employed the same strategies notwithstanding their L1. More specifically, in terms of the reading speed, 70% of speakers of English as L1 and 71.4% of speakers of English as L2 indicated that they read the text at their normal pace. When focusing their attention while reading, 56% of the speakers of English as L1 focused on both word meaning and form and only 14% chose "other" criteria. Similarly, 39% of speakers of English as L2 focused on the word meaning and form, and 9.3% preferred "other" criteria. Finally, the attention span was combinations of 2-3 words for 46% of speakers of

English as L1 and 42.9% for speakers of English as L2, while only 8% of speakers of English as L1 and 7% of speakers of English as L2 read the text word-by-word.

Although participants implemented the same test taking strategies, it yet remains to be seen how effective these strategies were in terms of their relationship to the test scores in the groups of speakers of English as L1 and L2.

Step 3. Differences/similarities in the test-taking strategies, additional languages, age, and gender between the groups of speakers of English as L1 and L2. A two-way between-groups analysis of variance (ANOVA) was used to describe statistically significant differences in the English as L1 and L2 participants' choices of the independent categorical variables, including:

(a) Focus of attention while reading, i.e., criteria of "word choice" ("meaning", "form", "form and meaning", and "other");

(b) Reading speed (slow, normal, and fast);

(c) Attention span while reading ("word-by-word", "combinations of 2-3 words", "combinations of 5 and more words", and "whole sentences");

(d) Another language (additional language for speakers of English as L1; and L1 for speakers of English as L2).

As a result of the data analysis at Step 2**,** the following similarities between the groups of speakers of English as L1 and L2 were identified:

The focus of attention while reading. The analysis confirmed that the focus of attention while reading (criteria of "word choice") had a statistically significant effect of on test scores [$F(3, 85) = 2.85$, $p = .042$), and the effect size was moderate (eta-squared = .09). The results of Levene's test ($p = .487$) suggested moderate homogeneity of

variances, and therefore, the significance of the effect. This means that the way speakers of L1 and L2 read and processed the text impacted their test scores.

The descriptive statistics indicated that in both groups the largest number of participants –28 speakers of English as L1 (56%) and 17 speakers of English as L2 (39.5%)– focused on both word meaning and form when identifying false collocations. This result agrees with Lewis (2000) and Sinclair et al. (2004), who suggest that meaning and form are highly interconnected. However, the highest test scores were achieved by the few participants –7 speakers of English as L1 (14%) and 4 speakers of English as L2 (9%)– who focused on "other" criteria, that is, something other than word form or meaning; meanwhile, the lowest test scores were achieved by those participants –11 speakers of English as L1 (22%) and 10 speakers of English as L2 (23%)– who focused on word form only.

Post hoc comparisons using the Tukey HSD test suggested that the mean score for the "other" condition of choice ($M = 53.5$, $SD = 26.3$) was significantly different from the "word form" condition of choice ($M = 30.7$, $SD = 22.3$), $p = .008$. However, the "word meaning" condition of choice ($M = 35.4$, $SD = 19.4$) and the "both" (word meaning and form) condition of choice ($M= 41.8$, $SD = 20.1$) did not significantly differ from the "other" and "word form" conditions. These results might indicate that the "other" criteria (described by the participants as "flow", "internal feeling", and "sentence structure") might influence the test scores in both groups of speakers (see Figure 1). These findings might be connected to the concept of intuition as discussed in Siyanova's (2010) and Siyanova-Chanturia and Martinez's (2014) research; however, more research is needed because very few participants (7 speakers of English as L1 and 4 speakers of English as

L2) selected this option. Additionally, such researchers as Macis and Schmitt (2017),

McGee (2006), Siyanova-Chanturia and Spina (2015), and Wray (2008) also notice that

intuition is somewhat subjective and might not work even in case of L1 speakers.
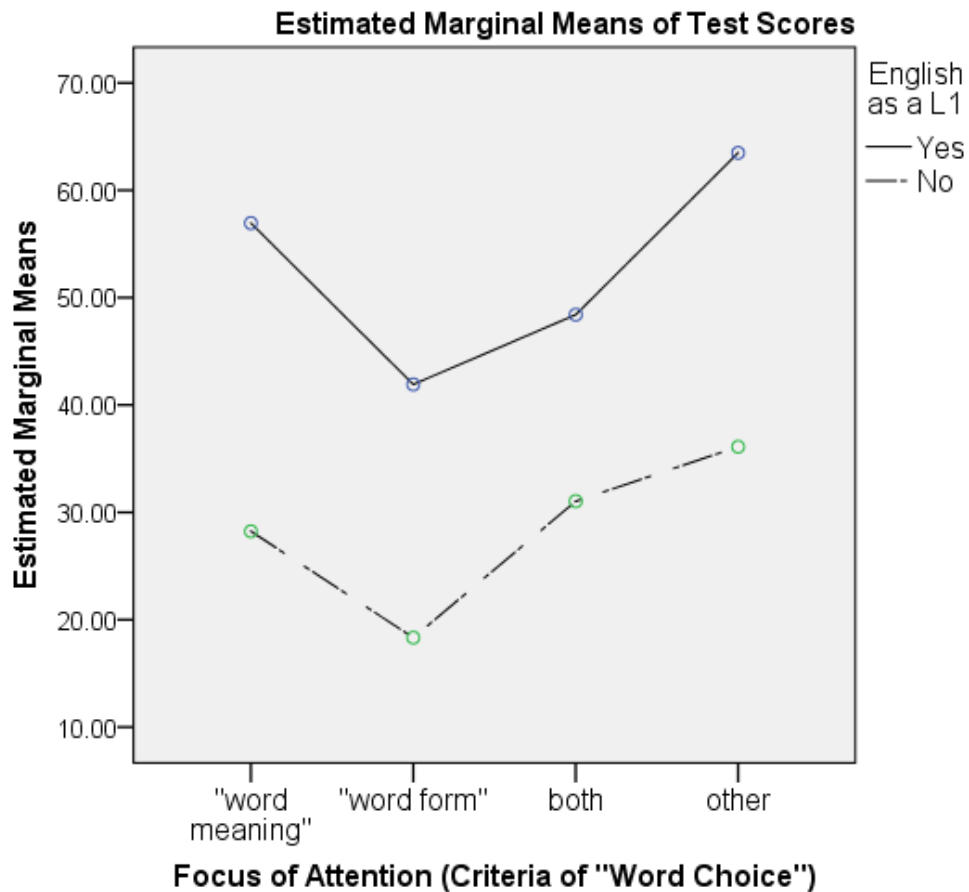
Figure 1. Means of test scores and focus of attention (criteria of "word choice") for

speakers of English as L1 and L2.

Reading speed. In the group of speakers of English as L1, descriptive statistics

showed that the 10 respondents (20%) who identified their reading speed as "slower than

usual" and the 5 participants (10%) who indicated their reading speed was "faster than

usual" scored higher on the test than the 35 participants (70%) who read the text at their

normal pace. On the contrary, in the group of speakers of English as L2, test scores were higher for the 30 participants (69.7%) who read at their normal pace, but lower for the 7 respondents who read more slowly (16.2%) and the 6 participants (13.9%) who read faster. However, the Sig. value ($p = .011$) in Levene's test was less than the minimum of .05 and indicated no homogeneity of variances. Additionally, the two-way between groups ANOVA showed that the differences in the reading speed did not interact with the test scores and were not statistically significant [$F(2,86) = 4.02, p = .600$]. The eta-squared statistic of .08 indicated a moderate effect size in terms of the mean values for the reading speed for speakers of English as L1 and L2.

These results suggest that although speakers of English as L1 and L2 differed in how fast they read, this fluency did not make a statistically significant contribution in terms of their test scores. This contradicts multiple studies by Henriksen (2013), Nation and Webb (2011), and Wang and Good (2007), who consider fluency an important part of collocational competence. One possible explanation might be that the format of self-assessment implemented in this study might not be an accurate tool for measuring reading fluency.

Attention span while reading. The descriptive statistics showed that in both groups (speakers of English as L1 and L2), "combinations of 2-3 words" was the most popular choice, selected by 23 speakers of English as L1 (46%) and 18 speakers of English as L2 (41.8%). Reading "word-by-word" was the least popular strategy, selected by 4 speakers of English as L1 (8%) and 3 speakers of English as L2 (6.9%). This indicates that English as L2 speakers are aware of the necessity to read in chunks. However, the results of the two-way between-groups ANOVA showed that these

differences were not significant in terms of test scores because there was no statistically significant effect of attention span on the test scores for the four conditions (reading the text word-by-word, in combinations of 2-3 words, in combinations of 5 and more words, or sentences) [$F(3, 84) = 1.16$, $p = .330$], and no interaction effect between the factor of English as L1 and the factor of attention span [$F(3, 84) = .56$, $p = .640$]. Therefore, although participants were guided by different criteria when reading, their preference of "word-by-word", "word combinations", or "whole sentences" reading strategies did not appear to significantly influence test scores. Just as it was the case with the previous variable of reading fluency, the validity of this finding is somewhat questionable due to the self-reporting format of the survey.

Another language (see Figure 2). For speakers of English as L1, the two-way between-groups ANOVA did not identify any statistically significant difference in the test scores between those who had some knowledge of an additional language (Arabic, Chinese, French, or other) and those who did not know any other language [$F(5,44) = .95$, $p = .460$], even though Levene's test indicated a large Sig. value of .800 and confirmed homogeneity of variances. This is a predictable finding, since their English is more developed and likely more influential in comparison to any additional languages they might know.

For speakers of English as L2, there was a statistically significant difference in test scores for different primary language groups [$F(3,39) = 5.41$, $p = .003$], which is confirmed by the results of Levene's test ($p = .810$) that suggested homogeneity of variances. The difference in the mean scores between the groups of speakers of Arabic, Chinese, French, and other languages was large (eta-squared = .29). Post hoc

comparisons using the Tukey HSD test showed that the mean scores for 17 speakers of Chinese as L1 (39.5%) were significantly lower ($M = 21.2$, $SD = 14.5$, $p = .006$) than the mean scores of 11 speakers (25.5%) of other languages as L1, such as French, German, Hindi, Portuguese, and Urdu ($M = 41.4$, $SD = 15.6$, $p = .006$). The same can be stated about speakers of Arabic as L1 (11 participants, 25.5%) ($M = 21.2$; $SD = 13.8$, $p = .013$) and speakers of additional primary languages. Since those participants whose primary languages were Arabic or Chinese scored lower on the test than speakers of such languages as French, German, Hindi, Portuguese, and Urdu, this finding suggests that L1 might impact recognition of false collocation in L2, which corresponds to existing research by Peters, (2016), Wolter and Gyllstad (2011, 2013), Wolter and Yamashita (2015), and Yamashita and Jiang (2010). However, unequal distribution of the language data (the majority of participants were Arabic and Chinese speakers) and the fact that Arabic and Chinese belong to a different language branch than Indo-European languages could potentially undermine the importance and generalizability of this finding.

Since it is not recommended to run ANOVA with continuous independent variables, a between-groups Pearson product-moment correlation coefficient (PPMCC) analysis was implemented to identify how participants with different ages and genders recognized collocations.

Age. The PPMCC analysis showed that there was no statistically significant relation between the age of participants and test scores in the group of speakers of English as L1 ($r = .07$, $p = .590$) or L2 ($r = .041$, $p = .790$). This might be due to the fact that most participants were approximately in the same age group of 18-19 - 40 (80%)

speakers of English as L1 and 22 (51%) speakers of English as L2 - and so there was not enough age diversity.

Gender. The PPMCC analysis showed that there was no statistically significant relation between the gender of participants and test scores in the group of speakers of English as L1 ($r = -.04$, $p = .790$) or L2 ($r = .03$, $p = .830$). This finding might be potentially mitigated by the fact that there was an unequal number of males (71%) and females (29%) participating in the study.
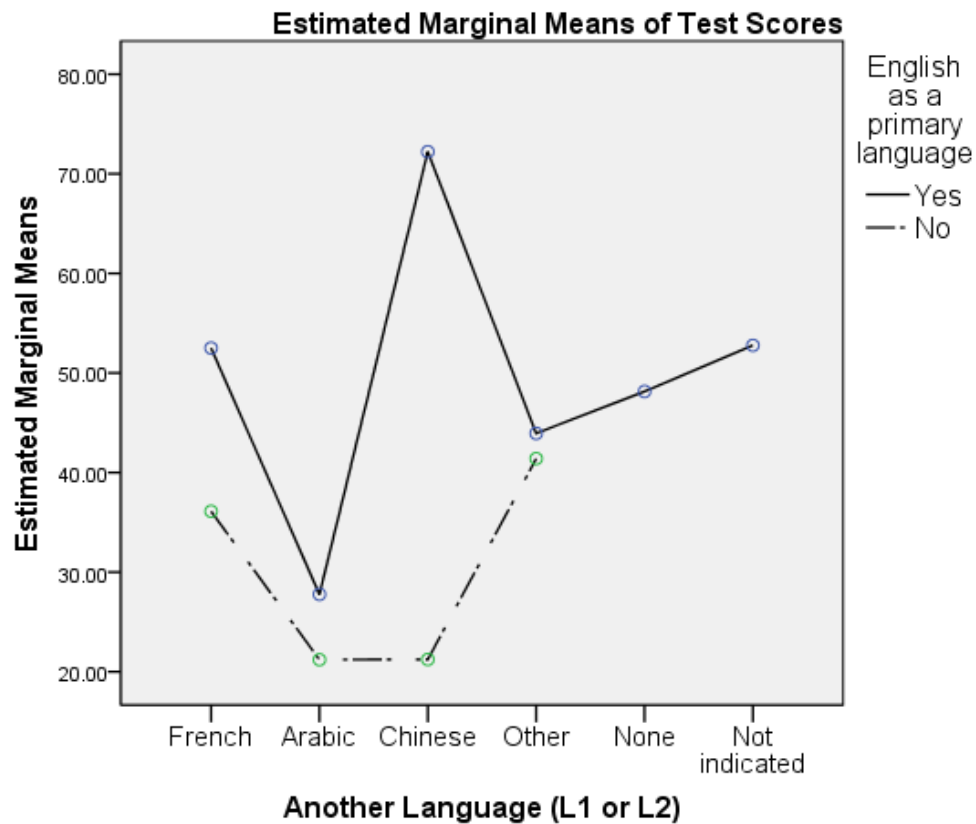
Figure 2. Means of test scores and another language for speakers of English as L1 and L2.

In sum, it appears that there are some differences and similarities between the groups of speakers of English as L1 and L2 in terms of the cognitive strategies used for approaching the test. Differences concern reading speed: speakers of English as L1 obtained higher test scores when reading either slower or more quickly than normal, and yet the opposite was the case for speakers of English as L2, who obtained higher test scores when reading at their normal pace. Attention span in both groups was similar: participants preferred reading chunks of text versus word-by-word or whole sentences. However, the factors of reading speed and attention span did not emerge as statistically significant. The most likely reasons for this are the lack of homogeneity of variances within each of the variables and the self-reporting format of the post-test survey. The external factors of age and gender also were not significant, possibly for the same reason of insufficient data variability (most participants were in the same age group of 18-19 and over 70% of the participants were male). The factor of another language was statistically important for speakers of English as L2 only, where their primary language predictably dominated their competence in English.

The focus of attention while reading (criteria of "word choice") emerged as a significant factor. Contrary to Lewis (2000) and Sinclair et al. (2004), it was neither word meaning nor form, but rather something different, defined by the respondents as "flow", "internal feeling", or "sentence structure" and potentially connected to intuition. However, this finding is somewhat questionable because only a small number of participants in both groups (9 people, that is, less than 15%) selected this option.

Step 4. Identifying potential factors of influence in the group of speakers of English as L2 only. A one-way ANOVA was run for the group of speakers of English as

115

L2 only to explore statistically significant differences among additional independent

categorical variables, which were specific to speakers of English as L2 only:

(a) Predominant language of communication;

(b) Vocabulary learning strategies ("individual unrelated words", "in

sentences/combinations", and "other");

(c) Self-assessment of the English language proficiency in comparison to their L1.

Next, a PPMCC analysis explored statistically significant differences in terms of

such independent continuous variables as self-assessment of English language

proficiency, age of onset of learning English, and the number of years participants have

lived in an English-speaking country.

Predominant language of communication (see Figure 3). The descriptive statistics

showed that 17 participants (39.5%) predominantly communicated in their primary

language, while 16 participants (37%) used both languages, and only 9 participants

(20.9%) used English only (potentially advanced bilinguals). A one-way ANOVA

analysis indicated that there was a significant effect of the predominant language of

communication on test scores [$F(2, 39) = 5.57$, $p = .007$], and the Levene's test statistics

($p = .980$) confirmed homogeneity of variances. Post hoc comparisons using the Tukey

HSD test identified that the largest mean score for the "English as a predominant

language of communication" option ($M = 43.2$, $SD = 14.9$) was significantly different

from the lowest mean for the "native language for communication" option ($M = 23.5$, SD

$= 15$, $p = .009$) and the mix of both languages option ($M = 24.64$, $SD = 15.7$, $p = .016$),

where the difference in mean scores between these three groups was large (eta-squared =

.22). The conditions of using predominantly the L1 or a mix of the L1 and English did

not seem to significantly differ ($p = .976$). Based on this, it is possible to conclude that the amount of communication in English positively impacts test scores, which corresponds to the findings of Gabrýs-Barker (2006), Wang and Shih (2011), and Xu (2015).
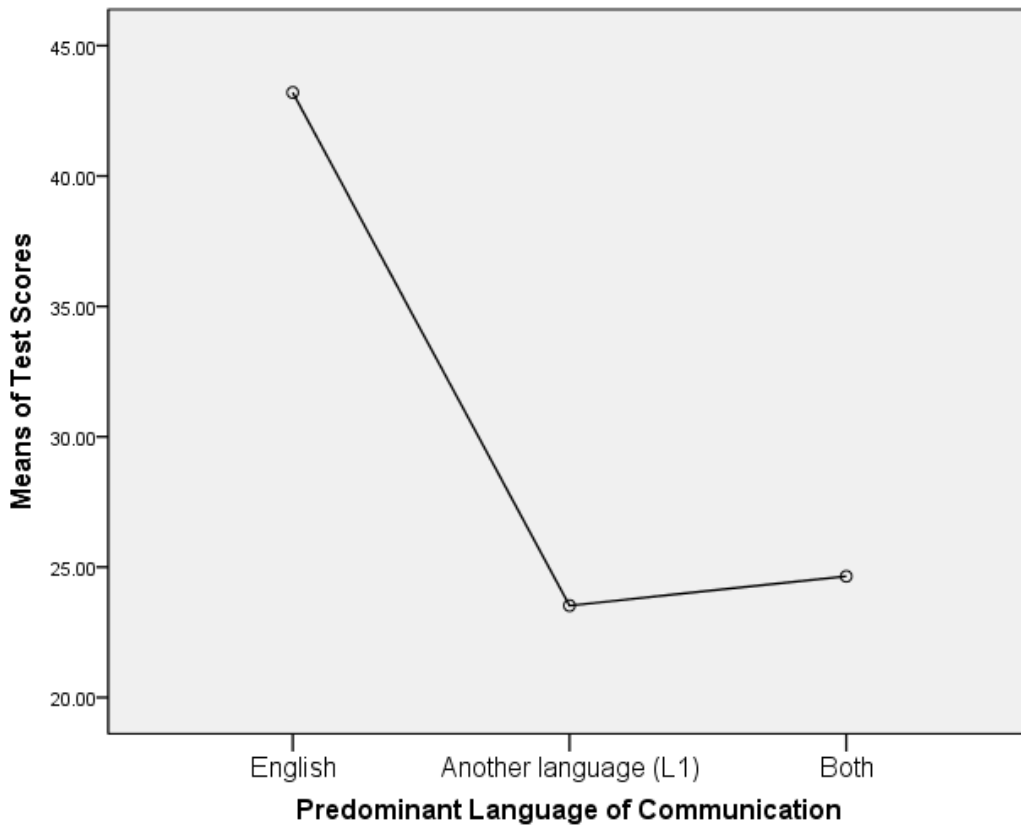
Figure 3. Means of test scores and predominant language of communication for speakers of English as L2.

Vocabulary learning strategies. First, a one-way ANOVA examined the differences in total test scores as potentially triggered by participants' preference for certain word learning strategies ("individual unrelated words", in

"sentences/combinations", and "other strategies"). However, there was no significant difference between word learning strategies in terms of test scores [F(2, 37) = 1.66, p = .444]. At the same time, the mean of the scores was the largest (M = 36.1, SD = 27.49) in the case of the "other strategies" choice, which only 2 participants (less than 5%) selected. These mean scores were closely followed by the "in sentences/combinations with other words" option (M = 29.83, SD = 17.23) that was preferred by 27 people (62.7%).

Due to the unequal sampling, the "other strategies" option was taken out, and an independent samples t-test focused on the "individual unrelated words" and "words in sentences/combinations" options only. There was a significant difference in the scores for "in sentences/combinations" (M = 29.83, SD = 17.23) and "individual words' (M = 19.99, SD = 9.51) vocabulary learning strategies [t(29.23) = -2.1, p = .036]. Therefore, learning words in combinations/sentences potentially has an effect on the total test scores, which agrees with studies by Lewis (2000), Nation and Newton (1997), and Schmitt (2000), who were the initiators of the explicit lexical approach to learning collocations.

English language proficiency in comparison to L1 (self-assessment). The results of an independent samples t-test showed that there was a statistically significant difference in the scores of those 16 respondents (37%) who indicated that their English language proficiency is "as good as their primary language proficiency" (M = 36.45, SD = 17.32) and those 26 respondents (60%) who considered their English language proficiency "worse than their primary language proficiency" (M = 23.07, SD = 14.71), [t (40) = 2.67, p = .011]. Additionally, Levene's test showed that the Sig. value was .234,

which means that variability was similar for both options. It is possible to deduce that an advanced level of proficiency in English does have an effect on improving test scores.

   English language proficiency level (self-assessment) (See Figure 4). Participants rated their English language proficiency level using a Likert scale from 1 ("unsatisfactory") to 5 ("excellent"). This rating was correlated with test scores. The PPMCC analysis showed that there was a statistically significant relation between how participants evaluated their proficiency level and their test scores ($r = .38$, $p = .010$), which means that their self-assessment positively correlated with their test scores.
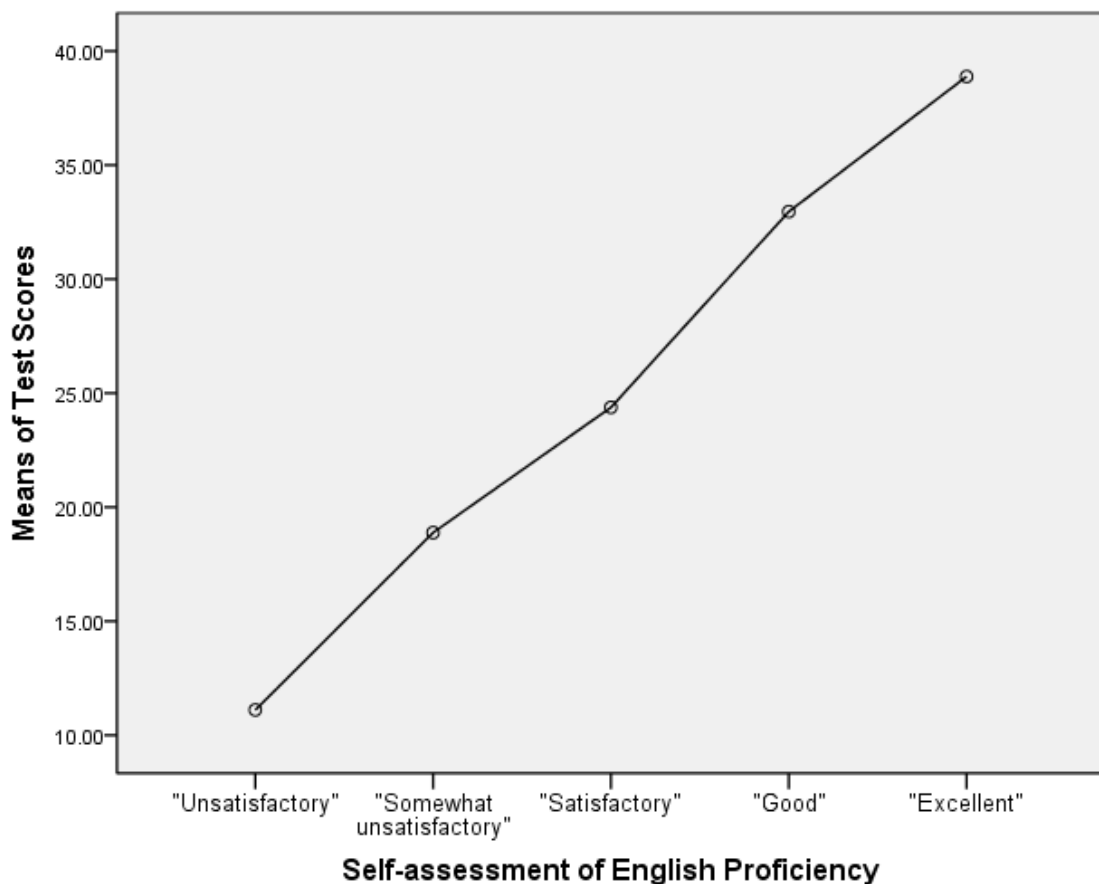
Figure 4. Means of test scores and self-assessment of English proficiency for speakers of English as L2.

Although some research (Edele, Seuring, Kristen, & Stanat, 2015) is skeptical towards self-assessment as a valid measurement, the results of this stage of the analysis indicate that ESL learners might have provided valid and accurate judgements of their language proficiency as it is related to their collocational competence, or more specifically, to their ability to recognize false collocations. Guduru (2014) likewise suggests that there is a relationship between vocabulary learning skills and self-assessment.

The length of residence in an English-speaking country. The PPMCC analysis suggested a moderate positive correlation ($r = .32$, $p = .043$) between total test scores and the number of years/months spent in an English-speaking country. This corresponds to the studies by Foster (2009) and Groom (2009) that suggest that collocational competency improves in an English-speaking environment. Since most participants (79%) spent less than five years in an English-speaking country, the evidence suggests that this correlation might have been even stronger if they had spent more time in an English-speaking environment. At the same time, this moderate correlation might suggest that other factors, such as the quality and quantity of input and output, are at play and can interfere with the effectiveness of the length of residence, as suggested by Adolphs and Durow (2004), Bardovi-Harlig and Bastos (2011), and Erman et al. (2015).

Age of onset of learning English. The PPMCC analysis showed that there was no statistically significant relationship between test scores and the initial age of learning English ($r = -.161$, $p = .321$), which might be due to the lack of data diversity: most

participants (62%) started learning English at approximately the same age (between 5 and 12 years old).

Overall, at Step 4 of the data analysis, several factors demonstrated statistically significant differences as related to test scores in the group of speakers of English as L2. First, potentially advanced bilinguals, that is, those participants who indicated their English language proficiency was as good as their L1 proficiency, had higher test scores. Second, English being participants' preferred language of communication positively impacted test scores in comparison to preferring their L1 as a primary language of communication; however, there was no statistically significant evidence that supported the idea of a mix/interchange/code-switching of two languages having an influence on the test scores, which somewhat contradicts the findings of Bylund et al. (2012), Liao (2010), and Millar (2011). One possible explanation might be the self-reporting format of the survey, which did not allow for a more accurate measurement of the amount of code-switching. Next, participants' self-assessment of their English language proficiency level generally corresponded to their test scores. Finally, learning words in combinations/sentences has a statistically significant advantage over learning individual unrelated words, which agrees with findings by Gitsaki (1999), Lewis (2000), and Wray (2002). Insufficient data variability might have interfered with the significance of the difference in test scores as correlated to the age of onset of learning English (emerged as non-significant) and the length of residence in an English-speaking country (emerged as only moderately significant).

The above analysis helped to examine and describe each of the factors separately, and how they compare across the two groups of participants. The next step was to compare these factors in their relation to each other.

Step 5. Comparing/correlating factors influencing recognition of false collocations in the groups of speakers of English as L1 and L2. The two groups of speakers of English as L1 and L2 were analyzed separately, where in each of the groups a stepwise regression explored the relationship between overall test scores and independent categorical and continuous variables. Since categorical variables are not typically used in regression, they were re-coded as "dummy" binary variables, in which 0 corresponded to the absence of a condition, and 1 equaled the presence of the condition, which made it possible to conduct the regression analysis.

Factors in the group of speakers of English as L1. The independent variables for the group of speakers of English as L1 were:

(a) Age;

(b) Gender;

(c) Reading speed (slow, normal, or fast);

(d) Focus of attention while reading - criteria of "word choice" ("meaning", "form", "form and meaning", and 'other"); and

(e) Attention span while reading ("word-by-word", "combinations of 2-3 words", "combinations of 5 and more words", and "whole sentences").

First, following Cohen (1988), a medium strength positive correlation was identified between the test scores and the independent variable of the focus of attention on "other" while reading ($r = .27$, $p = .030$), which means that those participants who

focused on something other than word meaning and/or form while reading obtained higher test scores. Additionally, a medium strength negative correlation was identified between test scores and the factor of "normal reading speed" ($r = -.34$, $p = .007$), which indicates that those participants who read the text at their normal pace scored lower than those participants who attempted to read more or less quickly than usual.

The results of the stepwise regression, summarized in Table 8, showed that when compared, neither of the factors except the "normal reading speed" ($r = -.45$, $p = .001$) and the focus of attention on something other than the word meaning and/or form (that is, the "other" option, defined by the participants as "general flow", "inner feeling", and "sentence structure") ($r = .38$, $p = .005$), emerged as strong score predictors since their beta values for standardized coefficients were below the minimum threshold of .3, and their Sig. values were more than .05.

The Sig. value of .001 in the ANOVA test box confirmed the large statistical significance of the analysis results. The model summary indicated the R-squared value of .257, which means that the model explains 25.7% of the variance in the test scores. The adjusted R-squared values demonstrated that adding the factor of "Criteria of word choice, option - other" to the factor of "normal reading speed' changed the percentage of predictability from 10 to 22.5; however, the overall percentage still remained small. Small VIF values (less than 10) and large tolerance values (more than .10) suggested no multicollinearity or singularity. Consequently, the normal probability plot (P-P) of the regression standardized residual chart (see Figure 5) demonstrated no significant deviations of the observed value from the expected value, which indicates data normality and linearity.

123

These results indicate that the relationship between the test scores and the normal

reading speed in the group of speakers of English as L1 appeared to be inverse, which

means that those participants who attempted to read at their normal speed actually had

lower test scores. Unlike the ANOVA results summarized at Step 3 of the data analysis,

the regression analysis showed that neither "slow reading speed" nor "fast reading speed"

were statistically significant score predictors. Consequently, we cannot come to any

conclusive results as to whether the speed of reading actually impacts test scores for

speakers of English as L1. The factor of the focus of attention on "general flow", "inner

feeling", and "sentence structure" seemed to have an impact on test scores in the group of

speakers of English as L1; however, according to the frequency counts, only 7

participants (14%) selected this option over the predominant majority of the participants

focusing on different criteria.

Table 8

Stepwise Regression for Speakers of English as L1. Independent Variables: "Normal

Reading Speed" and "Criteria of Word Choice, Option – Other"

Model		Standardized Coefficients		Collinearity Statistics	
		Beta	Sig.	Tolerance	VIF
1	(Constant)		.000		
	Normal Reading Speed	-.345	.014	1.000	1.000
2	(Constant)		.000		
	Normal Reading Speed	-.447	.001	.930	1.075
	Criteria of Word Choice_Other	.384	.005	.930	1.075

Normal P-P Plot of Regression Standardized Residual

Dependent Variable: Test Scores; Independent Variables: "Normal Reading Speed" and "Criteria of Word Choice, Option – Other"
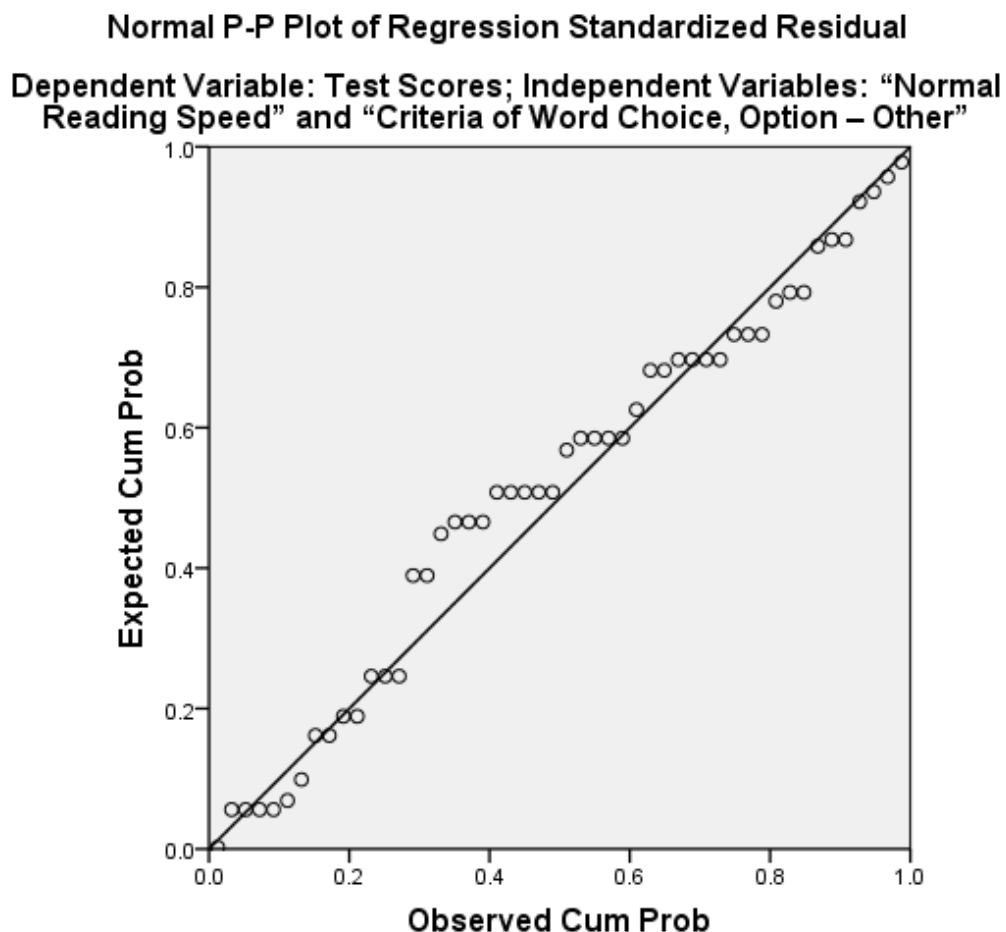
Figure 5. Normal P-P plot for speakers of English as L1. Independent variables: "Normal Reading Speed" and "Criteria of Word Choice, Option – Other".

Factors in the group of speakers of English as L2. The independent variables for the group of speakers of English as L2 were:

(a) Age;

(b) Gender;

(c) Reading speed (slow, normal, or fast);

(d) Focus of attention while reading - criteria of word choice ("meaning", "form", "form and meaning", and "other");

(e) Attention span while reading ("word-by-word", "combinations of 2-3 words", "combinations of 5 and more words", and "whole sentences");

(f) Predominant language of communication;

(g) Vocabulary learning strategies ("individual unrelated words", "in sentences/combinations", and "other");

(h) English language proficiency (self-assessment on the Likert scale);

(i) English language proficiency in comparison to L1 (self-assessment);

(j) Length of residence in an English-speaking country; and

(k) Age of onset of learning English.

First, a medium strength positive correlation was identified between test scores and the factors of English as the predominant language for communication ($r = .47$, $p = .001$), English proficiency as compared to L1 proficiency ($r = .39$, $p = .005$), self-assessment of English proficiency on the Likert scale ($r = .38$, $p = .006$), and the length of residence in an English-speaking country ($r = .32$, $p = .021$). This means that all of these factors that had already been identified as positively influencing test scores at Step 4 of the analysis confirmed to be statistically significant when correlated all together.

At the same time, a medium strength negative correlation emerged for the factor of the focus of attention on word form while reading ($r = -.31$, $p = .020$), which means that those participants who focused on the word form scored lower than those who focused on the word meaning or something else.

126

As Table 9 demonstrates, when running a stepwise regression analysis with these variables the only statistically significant score predictors were the factors of "word form" as the focus of attention while reading ($r = -.39$, $p = .006$), which emerged as a negative score predictor, and the factor of "English as a predominant language for communication" ($r = .53$, $p = .000$), which was a positive score predictor. The beta standardized coefficients were more than .30, and small Sig. values suggested the statistical importance of findings. The overall Sig. value of the model was .000, which indicated a large statistical significance; the R-squared value was slightly larger than in the previous analysis (.371), which means that the model explains 37.1% of the test score variance. The adjusted R-squared values showed that when added to the factor of "English as a predominant language for communication", the focus on word form while reading changed the original percentage of variance from 19.9% to 33.3%. The tolerance coefficients (more than .10) and VIF values (less than 10) were within the norm, which shows no multicollinearity or singularity. Consequently, the normal probability plot (see Figure 6) demonstrated no significant deviations of the observed value from the expected value, which shows data normality and linearity.

These results suggest that using English as the predominant language for communication might positively impact recognition of false collocations by speakers of English as L2. This corresponds to findings by Gabrýs-Barker (2006), Wang and Shih (2011), and Xu (2015), namely, that more advanced bilinguals for whom English takes over as their primary language might have an advantage over emergent bilinguals. Another finding of Step 5 of the research is that focusing attention on the word form negatively impacts the process of recognition, which, consequently, confirms the point of

view of Lewis (2000) and other supporters of the lexical approach to collocational

learning who downplay the word form in favour of meaning and context.

Table 9

Stepwise Regression for Speakers of English as L2. Independent Variables: "English as a

Predominant Language of Communication" and "Criteria of Word Choice, Option –

Form"

		Standardized Coefficients		Collinearity Statistics	
Model		Beta	Sig.	Tolerance	VIF
1	(Constant)		.000		
	Language of Communication_English	.471	.004	1.000	1.000
2	(Constant)		.000		
	Language of Communication_English	.530	.001	.977	1.023
	Criteria of Word Choice_Form	-.392	.008	.977	1.023

Normal P-P Plot of Regression Standardized Residual

Dependent Variable: Test Scores; Independent Variables: "English as a Predominant Language of Communication" and "Criteria of Word Choice, Option – Form"
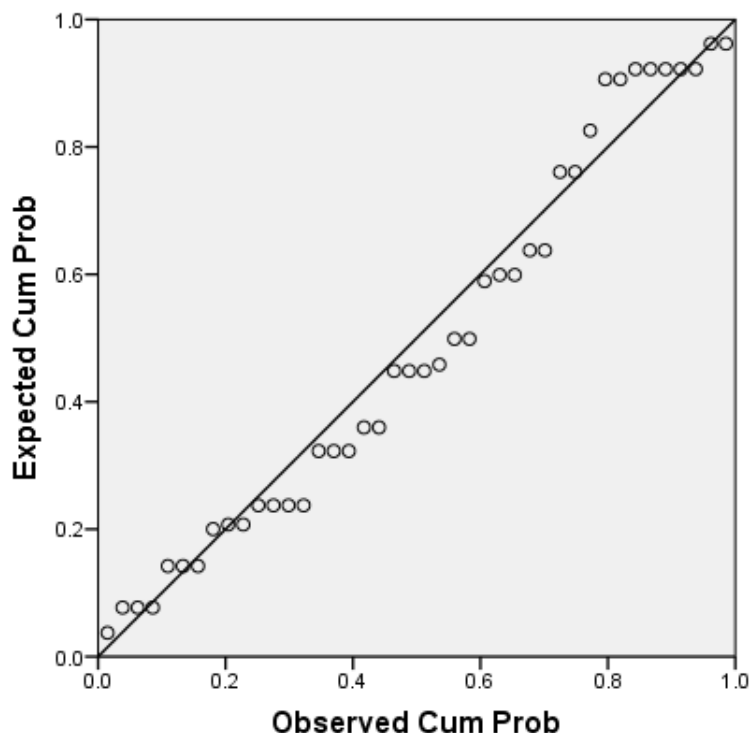
Figure 6. Normal P-P plot for speakers of English as L2. Independent variables: "English as a Predominant Language of Communication" and "Criteria of Word Choice, Option – Form".

Discussion. This large scale quantitative study, which included an acceptability judgement task and a post-test survey, was designed to explore and compare those external (collocation-non-specific) factors impacting recognition of false collocations by speakers of English as L1 and L2 that were identified in the pilot study as potentially influential. Other researchers have already explored most of these factors; yet, to the best

129

of my knowledge, they have combined and compared the factors into one comprehensive classification that incorporates:

1. Extralinguistic factors, such as age, gender, cognitive factors of vocabulary-learning strategies (Gitsaki, 1996; Lewis, 2000; Wray, 2002), reading fluency (Sonbul, 2015), the focus of attention and attention span while reading; and for the speakers of English as L2 only, age of onset of learning English (Granena & Long, 2013; Wray, 2002, 2008), length of residence in an English-speaking country (Foster, 2009; Groom, 2009; Lahmann et al., 2016; Schmitt, 2000), and language input (Cieślicka, 2015; Szudarski & Carter, 2016); and

2. Interlinguistic factors, such as English as L1 or L2, English proficiency level (Groom, 2009; Li & Schmitt, 2010), predominant language of communication (Wang & Shih, 2011; Xu, 2015), and L1-L2 interference (Irujo, 1986; Liao, 2010; Millar, 2011; Bylund et al., 2012).

The pilot study touched upon psychological-affective factors (e.g., motivation and attitudes toward language learning); however, due to time constraints, the focus of the research question, and subjectivity of the concepts of motivation and attitudes, these factors were excluded from the large-scale quantitative study. The pilot study additionally discussed the importance of criteria for recognizing false collocations that emerged as significant in other studies, namely, context and prior knowledge (Fan, 2009; McCarthy & O'Dell, 2005; Schmitt, 1997), style and genre (Lewis, 2000), intuition (Siyanova & Schmitt, 2008; Siyanova-Chanturia & Spina, 2015; Schmitt, 2010), analysis (Lewis, 2000), and language comparison (Kuo, 2009). However, these factors are highly variable and specific to participants' background, learning styles, and habits, and as such are not

easily measurable. Ideally, each of these would be further explored in a sub-study. These factors were also not considered at this particular research stage.

The quantitative SPSS-based analysis first examined potentially significant factors one-by-one in their relation to test scores and then compared them, which led to several interesting findings, most of which confirmed the findings of the studies discussed in Chapters 3 and 4.

First, the data analysis identified a statistically significant difference between test scores of speakers of English as L1 and L2. Predictably, speakers of English as L1 were able to identify twice as many incorrect collocations; thus, the interlinguistic factor of English as a primary language was the major factor impacting test scores, as has been suggested in other studies, including Siyanova and Schmitt (2008), Wolter and Gyllstad (2011, 2013), Wolter and Yamashita (2015), and Yamashita and Jiang (2010). This finding is further confirmed by another factor, English as the predominant language for communication (Wang & Shih, 2011), which emerged as a statistically significant positive score predictor in the group of speakers of English as L2 as contrasted to their L1 being predominantly used for communication. This result indicates the importance of the extralinguistic factor of being in an English language-speaking environment (language immersion) for mastering collocations, a finding that has been widely explored by supporters of the implicit teaching approach (Durrant & Schmitt, 2010; Pellicer-Sánchez, 2017).

Another extralinguistic factor connected to the importance of the English language-speaking environment is the length of residence (living/working/studying) in an English-speaking country. A moderately positive correlation between this factor and the

test scores was identified. However, it is important to note that most participants have lived less than five years in an English-speaking country while, according to Alsakran (2011), Foster et al. (2014), Groom (2009), and Wolter and Gyllstad (2013), a longer time span is needed for adult ESL learners to improve collocational competence.

Second, the analysis suggested that the ESL speakers' L1 might impact their recognition of false collocations because according to the findings, speakers of Arabic and Chinese as L1 scored lower on the test than speakers of other languages. This finding agrees with a number of studies, e.g., by Carrol et al. (2016), Peters (2016), Wolter and Gyllstad (2011, 2013), Wolter and Yamashita (2015), and Yamashita and Jiang (2010), which argue for L1 interference as the factor leading to overuse, underuse, or word-for-word translation of collocations, and therefore, collocational errors. However, a general lack of data diversity (most participants were speakers of Arabic and Chinese with only 25.5% of the participants speaking other languages) might undermine the importance of this finding.

The mix of both languages (L1 and English) when communicating, that is code-switching, which had previously been declared as important by Irujo (1986), Liao (2010), and Millar (2011), did not have a significant impact on test scores. However, the amount of code-switching and its impact on test scores is difficult to measure using a collocation recognition-based test and a self-reporting survey format, and more varied measurements, such as tests of productive collocational knowledge while writing and speaking and follow-up interviews, might prove useful in future research.

Another important finding is that speakers of English as L1 and those speakers of English as L2 who obtained higher test scores, that is, more advanced bilinguals,

132

exhibited some similarities in terms of how they approached test tasks. For instance, while reading the text and identifying non-English-like word combinations, they focused on word meaning and form combined, "general flow", "inner feeling", or the overall "sentence structure". On the contrary, those speakers of English as L2 who focused on the word form obtained lower test scores, and, as per a stepwise regression analysis, this factor was a statistically significant score predictor. This finding confirms that rather than focusing on grammar and structure as suggested by formalists (Stefanowitch & Gries, 2003), ESL learners need to treat form and meaning as interconnected parts of a whole, prioritize context, and develop their intuitive "feeling" of the language, which, according to Forsberg and Fant (2010), Moon (1998), and Syianova (2010), will lead them to perceiving word combinations holistically.

Some factors that were confirmed as significant in existing research on the ESL learners, that is, reading fluency and attention span while reading, did not appear significant in this study. This might be largely due to the significant limitations of the measurement format of the self-report and self-assessment survey, which, according to Edele et al. (2015) and Sitzmann, Ely, Brown, and Bauer (2010), might not be effective when measuring reading strategies due to potential biases and subjectivity. For future research, it might be useful to implement reaction time tests with a focus on reading errors or pauses as discussed in Rubin (2013), or short-term memory span tests (Jackson & McClelland, 1979).

Despite these results not being statistically significant, they still identified similarities between the ways speakers of English as L1 and L2 approached the task. For example, both groups of participants had a similar attention span (the majority preferred

133

reading the text in chunks). This means that speakers of English as L2 follow the same

reading strategy as speakers of English as L1 and are aware of the necessity to perceive

word combinations rather than focus on individual words. This might be the result of

their formal language instruction because, as literacy studies (Nation, 2009; Oliver &

Young, 2016) have argued, fast-paced reading and reading text in chunks might

contribute to fluency and better comprehension. Interestingly, the stepwise regression

analysis in the group of English as L1 speakers showed that those participants who

attempted to read at their normal pace scored lower on the test. This finding did not have

support in the literature; however, it might be explained by participants' unfamiliarity

with the task of proof-reading/revising, as the English as L1 speaker in the pilot study had

remarked. At the same time, as per the stepwise regression analysis in the group of

English as L2 speakers, the overall reading speed did not seem to have a statistically

significant impact on test scores, even though some differences were observed in terms of

higher scores for those who read at their normal pace and lower scores for those who read

more slowly or more quickly. This might be explained by either the task novelty, or, as

Siyanova-Chanturia and Spina (2015) and Schmitt (2010) suggested, the absence of a

well-developed intuition that would have helped them to identify incorrect collocations.

The usefulness of reading/perceiving text in chunks is further supported by the

results of a one-way ANOVA of the group of speakers of English as L2 that focused on

the factor of the preference for certain vocabulary learning strategies. Some studies

(Gitsaki, 1999; Lewis, 2000; McCarthy & O'Dell, 2005) argue that the lack of visible

progress in ESL students' learning of collocations might be explained by incorrect word

learning strategies (for example, learning individual unrelated words instead of word

combinations in their authentic context). The present study also identified a positive effect of learning words in "combinations/sentences"; however, the score variability as related to the choice of learning strategies was not the same as for other statistically significant variables, which makes the validity of this finding somewhat questionable.

Researchers (Foster et al., 2014; Granena & Long, 2013; Hyltenstam & Abrahamsson, 2003; Wray, 2002) have connected success in collocation acquisition with the age of onset of learning English; however, this study did not find a statistically significant correlation between test scores and the age at which participants began learning English. A possible explanation is that most participants started learning English between the ages of 5 and 12. However, as Granena and Long (2013) suggest, a younger age (0-6) might have benefitted learners because younger children are more adept at developing language intuition and holistic processing of word combinations. Another reason might be participants' background; they had most of their formal language education in a non-English-speaking country where they were not immersed in the target language environment.

The participants' age and gender also did not have an impact on test scores. While the fact that gender did not influence recognition of false collocations is expected, the findings about the age factor are worth further research. Studies by Han (2004), Schmitt (2010), and Wray (2002) have addressed vocabulary fossilization in adult learners; in contrast to these studies, and in agreement with Granena and Long's (2013) study, the results of the present research imply that collocation recognition is a life-long process and might not be prone to fossilization. However, it is also important to remember that this

finding might have been impacted by the lack of data diversity, as most participants were of approximately the same age (18-19).

An additional finding that might be of interest to educators, in particular those who debate the usefulness of peer-feedback and self-reflections, is a statistically significant positive relationship between participants' self-assessment of their English language proficiency level and overall test scores that emerged as a result of the one-way ANOVA of the group of speakers of English as L2. The post-test survey included two questions on English language proficiency. The first question asked participants to rate their knowledge of English on a Likert scale from 1 ("unsatisfactory") to 5 ("excellent"), and the second question required them to compare their L1 and L2 proficiency from "worse" to "the same". Those participants who indicated their English language proficiency was "as good as L1 proficiency" and rated their proficiency at 4-5 scored higher on the test. This finding confirms the results of studies by Guduru (2014) and Kayler and Weller (2007) that observed the benefits of self-assessment in the classroom, and shows that ESL learners might be able to accurately evaluate their progress in language learning and their language skills. This finding could be used in planning learning activities involving peer feedback and self-reflection.

Summary and Implications for Stage II

Chapter 5 focused on the two-stage study that explored the significance of the external (non-collocation-specific) extralinguistic and interlinguistic factors that might impact recognition of false (non-English-like) Adjective + Noun and Verb + Noun collocations and identified whether different or similar factors might be important in the case of speakers of English as L1 or L2. First, the pilot study included an acceptability

judgement task and a cloze/multiple-choice test, followed by a post-test interview with five participants (one speaker of English as L1, one advanced bilingual, and three emergent bilinguals). The results of the pilot study confirmed that the format of cloze/multiple-choice tests was inadequate for measuring recognition of whole collocations, which corresponds to Revier's (2009) findings. The results further identified some potentially significant factors and criteria that influence recognition of false collocations and determined that these factors might be similar, but not quite the same for speakers of English as L1 and L2. Moreover, the predominance of certain influential factors might vary depending on English level proficiency (emergent versus advanced bilinguals).

On the basis of this pilot study, a large-scale quantitative study was conducted with 50 speakers of English as L1 and 43 speakers of English as L2. The majority of ESL participants were in the same age group and shared the same L1 (Arabic or Chinese), age of onset of learning English, number of years of formal instruction, and length of residence in an English-speaking country, which might have limited data variability. This study implemented the same acceptability task format targeted at recognition of false collocations and a self-assessment-focused post-test survey based on the factors that appeared significant in the pilot study. Due to time constraints, there were no open-ended questions in the survey, and although participants were encouraged to comment on each question, multiple-choice and Likert scale questions were implemented, where the choice options were limited and pre-set. Additionally, some questions targeted reading and learning strategies, and language proficiency required self-assessment, which might not have been an optimal format. The lack of data diversity and the self-evaluative

137

component of the survey might have been a limitation to some extent. For example, such factors as the age of onset of learning English, reading fluency, and attention span (reading words one-by-one, in chunks, or whole sentences) were not identified as statistically significant in this study.

However, a number of valuable findings were identified regarding how speakers of English as L1 and L2 recognized false collocations and what factors contributed to this process. As a result, a typology of extralinguistic and interlinguistic factors influencing recognition of false collocations by speakers of English as L2 was developed. Among the primary factors of influence are English as L1, ESL speakers' L1, English as a predominant language for communication, vocabulary learning strategies (learning words in combinations), and to some extent, length of residence in an English-speaking country and participants' self-assessment of their English language proficiency. These results suggest that ESL learners might benefit from learning collocations rather than one-word lexical units, participating in language immersion-targeted programs and thus increasing the quantity and quality of language input and output, and using self-reflection and self-assessment to monitor their vocabulary learning progress.

Additionally, speakers of English as L1 and speakers of English as L2 who obtained higher test scores used similar reading strategies, that is, the focus on both word meaning and form, "general flow", "internal feeling", and "sentence structure" while reading the text and searching for incorrect word combinations. This result suggests that as English language proficiency and collocational competence grow, ESL speakers tend to successfully acquire word recognition/text processing strategies that are implemented by speakers of English as L1.

Stage I of the study focused on the external rather than internal (collocation-specific semantic and morphostructural) factors. Consequently, for the purpose of this stage, the test items –general English short-span 2-3-consecutive-words-long collocations– were selected at random, mostly on the basis of their morphosyntactic type (Adjective + Noun and Noun + Verb). It would be useful to explore how lexical and structural characteristics of collocations might impact their acquisition in an academic context. Thus, Chapters 6-8 discuss two large-scale quantitative studies that specifically focused on characteristics of academic English collocations and how these characteristics might influence collocation recognition and controlled production.

Chapter 6: Stage II. Internal (Collocation-specific) Factors

Rationale and Design

Stage II of the research focuses on the characteristics of collocation, namely the frequency of occurrence, the frequency of co-occurrence, and morphosyntactic type (MST), as three potentially significant internal factors influencing collocation recognition and controlled production. These factors have been widely explored and emerged as significant in the literature (Boers et al., 2014; Durrant & Schmitt, 2009, 2010; Granger & Bestgen, 2014; Li & Schmitt, 2010; Siyanova-Chanturia & Spina, 2015; Szudarski, 2016; Szudarski & Conklin, 2014; Wolter & Gyllstad, 2013). However, with only a few exceptions - such as Ellis et al.'s (2008) research that examined the length of collocational strings, the frequency of occurrence, and co-occurrence; Nguyen and Webb's (2016) study that explored the frequency of occurrence and co-occurrence and Adjective + Noun and Verb + Noun MSTs among other factors; and Peters' (2016) research that compared Adjective + Noun and Verb + Noun collocations - there have been no attempts to combine these three factors and compare them in relation to one another.

First, a collocation recognition and controlled production-targeted test was designed, modelled after Revier's (2009) CONTRIX test. The choice of test format was stipulated by the necessity for recognition and controlled production of the whole collocations and, therefore, is more applicable for the purpose of this study than other test formats, for example, standardized multiple-choice tests (Boers et al., 2014; Eyckmans, 2009; Gyllstad, 2009; Jafarpour et al., 2013; and Nguyen & Webb, 2016) that measure recognition and production of a part of collocation only (either a node or a collocate)

rather than the whole collocation. Second, this stage of the study focused on selecting test items on the basis of the above-mentioned three collocation-specific factors. These items formed a list of academic English collocations that are recommended for teaching in EAP programs and were used for the teaching and testing tasks in Stage III of the research. The focus on academic rather than general collocations is stipulated by the requirements of these programs and participants' (students') needs. Finally, a post-test survey was designed and administered for the purposes of collecting background information on the participants (demographics and language learning experience) and their perceptions of the test difficulty.

Instruments

Test format. The collocation recognition and controlled production-based test was modelled after the CONTRIX test (Revier, 2009) and consists of 34 items. Each item represents a sentence with a cloze gap retrieved from the academic section of the BNC and COCA. The content of the sentences does not pertain to one particular academic discipline; instead it is related to general aspects of student life, such as academic subjects and activities (going to the library, writing a report, working in groups, etc.). Consequently, the test does not require any knowledge of specialized academic terms and is focused on general academic English.

Participants were asked to fill in the gap with an appropriate word combination, that is, collocation. For this purpose, they referred to a two (or three)-word-choice and two (or three)-column matrix. Participants needed to choose a correct item for each of the two (or three) columns, so that all words together comprised one collocation appropriate for the cloze gap. According to Revier (2009), such a test format allows to measure the

141

productive knowledge of whole collocations rather than their individual components. However, unlike Revier's version that used three perfectly acceptable English word combinations (two distractors and one contextually appropriate collocation), the test in this study followed Jafarpour et al. (2013), Koya (2005), Nizonkiza (2015), and Webb and Kagimoto (2011) by using synonyms for collocational components as distractors. Thus, only one word combination could be both acceptable in English and appropriate in the given context. The rationale for this decision was the necessity for participants to be able to both recognize correct collocations and differentiate them from non-English-like word combinations (Barnbrook et al., 2013; Schmitt, 2010; Wray & Perkins, 2000), that is, engage in an acceptability judgement task. According to Wray (2008) and Webb et al. (2013), being able to provide acceptability judgement is an important aspect of collocational competence that signals the participants' knowledge of the form and meaning of collocation, fluency, and intuition.

Criteria for the selection of test items. While Revier's test focused on Verb + Noun collocations only with nodes (verbs) selected on the basis of their frequency and semantics (polysemy) from the pool of general English, the test in this study used different criteria for selecting test items. These criteria, summarized in Table 10, were determined based on collocation characteristics identified in existing research and examined in relation to one another.

Table 10

Criteria of Selecting Test Items

	Adjective + Noun	Verb + Noun	Verb + Adjective	Verb + Adverb	Adverb + Adjective	Total for all MSTs	TOTAL
High F - High MI	3 test items	3 test items	2 test items	2 test items	2 test items	12 test items	34 TEST ITEMS
Low F - Low MI	3 test items	3 test items	2 test items	0	2 test items	10 test items	
Low F - High MI	3 test items	3 test items	2 test items	2 test items	2 test items	12 test items	

Notes:

F – Frequency of occurrence

MI – Mutual information

MST – Morphosyntactic type

Frequency of occurrence (F). The majority of corpus-based collocation studies

focus on the frequency of occurrence of whole collocations in a corpus (Evert, 2008;

Granger & Bestgen, 2014; Koya, 2005; Siyanova-Chanturia & Spina, 2015; Sonbul,

2015; de Souza Hodne, 2009; Wolter & Gyllstad, 2013), with fewer studies examining

the frequency of individual components, such as node frequency (Durrant & Schmitt,

2010; Nizonkiza, 2015) and collocate frequency (Hoffmann & Lehmann, 2000). Since

this study measures participants' knowledge of whole collocations rather than their parts,

it focuses on whole collocations with low and high frequency of occurrence in the corpora.

However, certain incongruity in regards to the frequency threshold has been observed. For example, while Szudarski and Conklin (2014) set the threshold of 30 times per 100 million in the BNC and differentiate between infrequent collocations that occur fewer than 30 times and frequent collocations that occur 79 times and more, Wolter and Gyllstad (2013) suggest the minimum threshold of 10 times per million in the COCA; other studies provide yet different numbers depending on the size of the corpora, the length and structure of collocations, and data collection objectives. Consequently, rather than establishing an absolute threshold, this research followed Granger and Bestgen (2014) and Siyanova-Chanturia and Spina (2015) and identified low and high normed frequency bands on the scale of continuity and average frequency. For instance, while more numerous Verb + Noun collocation items had high frequency of 10-11 per million, less numerous Adverb + Adjective collocations had high frequency of only 5-8 per million. The average high frequency measured across the total of 12 test items with high frequency was 12.08. The average low frequency was 1.31 for the total of 22 items.

Frequency of co-occurrence (MI). The relations between the node and its collocate are manifested in the strength of association (Ellis et al., 2008; Gablasova et al., 2017; Nguyen &Webb, 2016) as measured by a low and high MI score, which served as the second influential factor in this study. According to Ackermann & Chen (2013), Barnbrook et al. (2013), Durrant & Schmitt (2009), Ellis et al. (2008), Fellbaum (2007), and Ngyuen & Webb (2016), the MI score indicates the probability that two words co-occur not by chance. It shows the difference between the observed and expected

frequency of co-occurrence and is calculated by identifying the proportion (percentage) of the number of times that words co-occur versus the number of times that words do not co-occur in the corpora. Consequently, the MI score identifies formulaicity of a word combination by determining the degree of meaningfulness of the semantic association between the words.

Following research by Ellis et al. (2008), Nguyen and Webb (2016), and Simpson-Vlach and Ellis (2010) who differentiated collocations with low and high MI score bands, no absolute maximum threshold was established. The average high MI in this study was 8.69 (for 24 items). However, following Hunston (2002) and Nguyen and Webb (2016), the minimum threshold for two-word-long collocations was set at 3. The average low MI score in this study was 3.99 (for 10 items).

Morphosyntactic type (MST). Another criterion was the morphosyntactic structure of a word combination. While most previous studies focus on one specific type of collocation, this study followed Ackermann and Chen (2013) and examined Adjective + Adverb, Adjective + Noun, Verb + Noun, Verb + Adjective, and Verb + Adverb collocations. According to Benson et al. (1997), Gitsaki (1996), and Poulsen (2005), the above-mentioned five morphosyntactic types are the most frequent and prototypical.

Some of these MSTs allow for variation in the syntactic position of their components. For instance, in the case of Verb + Adverb collocations, the adverb can either precede or follow a verb. F and MI counts might be slightly different for each case. For example, in the collocation *widely used*, where the adverb *widely* precedes the verb *used*, F is 18.49 and MI is 7.76. However, in the collocation *used widely,* in which the adverb follows the verb, F is 1.97 and MI is 4.53.

Additionally, variation occurs at the level of the word form. Thus, verbs can be in the form of an infinitive, gerund, or past participle, which in turn would influence the F and MI scores. For example, in the case of the aforementioned Adverb + Verb collocation *used widely,* the verb takes on the form of past participle (*used*), and F is 1.97 and MI is 4.53. However, in the case of the Adverb + Verb collocation *vary widely*, we have the verb in its infinitive form (*vary*), and F is 1.88 and MI is 8.39.

In order to eliminate the factor of variation, this study did not focus on the differences in the syntactic position or word forms. Whenever the difference in the F and MI scores was significant, as in the case with the collocation *widely used/used widely*, average F and MI scores were calculated.

Length of collocational strings and span. According to Simpson-Vlach and Ellis (2010) and Sinclair et al. (2004), collocations can range between two to five consecutive and/or non-consecutive words (including lexical/notional words, such as nouns, verbs, adverbs, and adjectives, and grammatical words, such as prepositions, auxiliaries, articles, etc.). Since a longer span could potentially influence collocation processing and thus become an additional distractor, this study focuses on short-span adjacent collocations (Vechtomova et al, 2003), that is, combinations of two consecutive lexical units which represent minimalistic collocations.

The 34 test items (see Appendix H) were manually selected from the Academic Collocation List (ACL) compiled by Ackermann and Chen (2013) on the basis of the Pearson International Corpus of Academic English (PICAE) that comprises academic texts from five English-speaking countries (Australia, Canada, New Zealand, UK, and the USA). This is the most comprehensive list of academic English collocations and consists

146

of 2,468 items belonging to different parts of speech. Permission was given by the authors (see Appendix I) to use their sets of initial data for manual vetting that contain information about the normed F per million and MI score.

Each test item represented one of the five MSTs and ranked from high to low on the scale of F and MI, thus combining three defining collocation characteristics. Since there was no absolute threshold for F and MI, the items were considered as oppositions of high F – low MI if their F score was significantly higher than their MI score, and conversely, low F – high MI if their MI score was higher than their F score. For the items labelled as high F – high MI, the score was expected to be at least 3 counts higher than the score of the items labelled as low F – low MI, and as approximate to the average high and low F and MI scores as possible. For example, in the case of Adverb + Adjective collocations, high F – high MI items were such collocations as *mutually exclusive* (F = 5.56, MI = 13.24) and *slightly different* (F = 9.6, MI = 7.35), and low F – low MI items were such collocations as *particularly appropriate* (F = 1.57, MI = 4.59) and *relatively common* (F = 1.17, MI = 4.25).

Difficulties in the item selection process were related to identifying a more or less equal number of collocations that would correspond to the above-mentioned criteria of the binary oppositions of low – high F, low – high MI, and five MSTs. The ACL has an unequal distribution of collocations belonging to different MSTs; thus, according to Ackermann and Chen (2013), over 74% (1,835 items) of the list is comprised of Adjective + Noun collocations, while Verb + Noun and Verb + Adjective collocations comprise only 13.8%, Adverb + Verb collocations 6.9%, and Adverb + Adjective 5% (p. 240). As a result, this study focused on 9 Adjective + Noun and 9 Verb + Noun

collocations with 3 items per each of the three categories (high F – high MI; low F – low MI; and low F – high MI); 6 Verb + Adjective and Adverb + Adjective collocations with 2 items per each of the above-mentioned three categories; and 4 Adverb + Verb collocations with 2 items per each of the two categories (High F – high MI; low F – high MI). No low F – low MI Adverb + Verb collocations were identified in the list.

The fourth category of high F – low MI collocations was excluded altogether because when analyzing items from the ACL, it became apparent that the lower the MI and the higher the F score, the less obvious ("weaker") collocations would be. For example, the word combination *other areas* has a very high F score (17.19) and a relatively low MI score (4). However, this word combination can hardly be considered a collocation since both words can easily combine with other words and are not strongly associated with one another. In comparison, the word combination *informed consent* has a lower F score (7.36) and a higher MI score (11.45), which makes it a more obvious ("stronger") collocation. These findings are in accordance with Edmonds and Gudmestad's (2014), Ellis et al.'s (2008), and Nguyen and Webb's (2016) studies that demonstrate that it is the frequency of co-occurrence rather than the frequency of occurrence that should serve as a defining characteristic and potentially the most influential factor.

Piloting the test.

Pilot study design. In order to ensure that the test items and format were clear and the model was a valid and reliable measurement of collocation recognition and controlled production, a two-stage exploratory pilot study was conducted. At the first stage, a draft version of the test was administered to two experts, both professors of Applied Linguistics disciplines who were L1 speakers of English. The experts were informed of

the study objectives, provided with the score of the test items on the scales of F and MI, and were asked to make an independent judgement of face validity of the items based on the following criteria:

1. Academic appropriateness of collocation and sentence items;

2. Contextual clarity of sentence items; and

3. Contextual appropriateness of the distractors.

At the second stage of piloting, the test was administered to and post-test semi-structured interviews were conducted with 14 participants, 7 in person and 7 online. These participants were intermediate to advanced ESL speakers from different primary language backgrounds (Arabic, Chinese, Korean, Portuguese, Romanian, and Russian) and age groups (19 to 40 years old). Participants were not given any explanations of the concept of collocation in order to ensure spontaneous recognition and production; however, they were provided with explicit instructions to complete the test without referring to any sources, such as dictionaries or grammar textbooks. They were not given a specific time limit for the test completion because at this stage it was important that they work at their own pace so that, based on the average time it would take them to complete the task, time limits for the actual test could be established. The average time it took participants to complete the test was 45 minutes.

Data analysis. Data from the pilot test were coded and analyzed using SPSS Version 22. When processing the pilot test results, an important decision had to be made, namely, how to enter and code data so as to identify different scales of F and MI throughout five MSTs, and what types of collocation errors would be considered for the analysis.

Coding collocational errors. As described in previous research (Benson et al., 1986; Hausmann, 1989; Sinclair et al., 2004), collocations consist of a minimum of two interdependent lexical units of which one is a node and the other a collocate. In addition, collocations might include grammatical units (prepositions and/or articles). Consequently, there can be several types of collocational errors:

1. Wrong node;

2. Wrong collocate;

3. Both the node and its collocate are wrong; and

4. Wrong grammatical unit(s).

While piloting the test, I considered these four types as separate categories of errors. However, the results demonstrated that participants selected correct lexical items of a collocation and yet scored low because of incorrectly selected articles and prepositions. There is no agreement in the research as to whether knowledge of grammatical units is a part of collocational competence. Some researchers (Benson et al., 1986; Gitsaki, 1996; Sinclair et al., 2004) argue that errors in choosing prepositions and articles should be considered collocational errors, while others (Hausmann, 1989; Nesselhauf, 2005; Wanner et al., 2006; Wanner, Ferraro, & Moreno, 2017) suggest that collocations consist of lexical units and thus they exclude grammatical fixed multiword combinations, such as Noun/Verb + Preposition, from their classifications of collocations. For these researchers, although function words might help to connect lexical units, errors when selecting a function word along with identifying a correct word form would be considered the domain of grammatical knowledge.

150

This study follows the latter perspective. While articles and prepositions were left in the matrix to ensure overall coherence, when analyzing the data it was decided to focus on lexical collocation components only and leave out grammatical errors (articles and prepositions). For example, for the collocations *play a role* or *raise awareness of*, the wrong choice of verb or noun was marked, but errors in selecting function words were not marked. Consequently, a test response was considered wrong only if there was an error in the choice of node, collocate, or both. Due to the large number and complexity of F, MI, and MST variables and their combinations, and in order to rule out possible variation and interpretation challenges by the participants, it was deemed not feasible to classify these three types of errors as separate variables. Thus, errors when choosing a node, a collocate, or both were all considered as one generic error.

To illustrate, when selecting words from the collocational matrix to fill in the blank in Sentence 14 (*"Universities are prohibited to _____ personal information about students"*), the following variations could occur: (1) supply access to (an error in the choice of node word); (2) give path to (an error in the choice of collocate); (3) supply path to (both lexical components are incorrect); (4) give access into (both lexical components have been selected correctly, yet the preposition is wrong); or (5) give access to (all components haven been identified correctly). Variations 1, 2, and 3 would all be considered as one generic error, while Variation 4 would not be considered as an error since the lexical components are correct. Variation 5 represents the correct choice.

Coding independent variables. Since some researchers (Durrant & Schmitt, 2009; Siyanova & Schmitt, 2008) consider F and MI as highly interconnected, and others (Ellis

et al., 2008; Ngyuen & Webb, 2016) examine them as separate variables, it was decided to consider these variables first as separate and then as combined entities.

Since it was necessary to identify which of the independent variables (frequency of occurrence, frequency of co-occurrence, morphosyntactic type, or combinations of these factors) might influence overall test scores (dependent variable), a matrix of variables was implemented (see Appendix J). In this matrix, the factors of five MSTs (different F and different MI), low F (different MI and different MSTs), and high MI (different F and different MSTs) were identified and singled out as seven main independent variables. However, it was not possible to single out high F (different MI and different MSTs) and low MI (different F and different MST) as two other separate independent factors, since according to my findings from the ACL (discussed in the section "Instruments. Criteria of Selection of the Test Items"), combinations of high F – low MI are unlikely because in these cases the word combination would not be fixed and would lose its collocational value. Consequently, while collocations with a low F score might have either a high or low MI score, collocations with high F might have a high MI score only. Similarly, collocations with a high MI score might have either a high or low F score, but collocations with a low MI score are most likely to have a low F score only. Different combinations of the above-mentioned seven main factors represented additional variables as summarized in Appendix H and Table 10.

The mean score of the pilot testing was 19.28 (56%) with the highest score of 23 and the lowest score of 14. To ensure test validity and internal consistency reliability, Cronbach's alpha test was run (Table 11). The alpha coefficient of .93 indicated high internal consistency of the test items, where according to Pallant (2010) and

152

Radhakrishna (2007) the alpha coefficient should be more than .70. Neither of the alpha values for "if item deleted" exceeded the original alpha coefficient value. In the corrected item-total correlation, values were more than .30, which means that each test item is highly correlated with the total test score.

After completing the test, participants were interviewed based on the following guiding topics:

1. Demographics (age, first language and any additional languages they might know, age of onset of learning English; how many years spent learning English);

2. Vocabulary-learning strategies; and

3. Test feedback:

   a. Which test items were the easiest/most difficult, and why?

   b. How would they rate the test in terms of its difficulty on a five-point Likert scale from 1 ("very easy") to 5 ("very difficult"), and why?

The first two questions targeted potentially influential extralinguistic and interlinguistic factors of L1 interference, the age of onset, length of formal and informal instruction, and input/learning strategies. The last question helped the researcher identify if/which test items might lack clarity or be difficult for the participants.

While the statistical analysis of the pilot test confirmed that independent variables accurately measured and were able to predict test scores, the analysis of qualitative interview data indicated that the test task and instructions needed to be simplified. Descriptive coding was conducted on the interview transcripts in order to identify common patterns that might help to modify test items. The emerging patterns are summarized in Table 12.

Table 11

Pilot Test Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Number of correct collocations	73.5000	713.838	.996	.937
Low F – Different MI – All MSTs	79.0769	808.981	.936	.931
Different F – High MI – All MSTs	77.9103	808.602	.931	.931
Adjective + Noun – Different F – Different MI	83.7564	938.810	.650	.935
Verb + Noun – Different F – Different MI	83.3077	912.761	.774	.933
Verb + Adjective – Different F – Different MI	84.9487	953.192	.590	.936
Verb + Adverb – Different F – Different MI	85.3462	956.801	.626	.936
Adv + Adjective – Different F – Different MI	85.1667	947.959	.585	.936
High F – High MI – All MSTs	81.7051	900.315	.768	.933
Low F – Low MI – All MSTs	82.9103	904.654	.788	.933
Low F – High MI – All MSTs	83.4487	905.471	.797	.933
Adjective + Noun – High F – High MI	86.0641	981.905	.383	.938

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Adjective + Noun – Low F – Low MI	85.7436	967.855	.552	.936
Adjective + Noun – Low F – High MI	86.4615	993.447	.206	.938
Adjective + Noun – Low F – Different MI	84.9487	959.010	.564	.936
Adjective + Noun – Different F – High MI	85.2692	973.030	.425	.937
Verb + Noun – High F – High MI	85.3077	972.917	.480	.937
Verb + Noun – Low F – Low MI	86.1282	976.451	.461	.937
Verb + Noun – Low F – High MI	86.3974	967.697	.655	.936
Verb + Noun – Low F – Different MI	85.2692	942.303	.676	.935
Verb + Noun – Different F – High MI	84.4487	936.952	.732	.934
Verb + Adjective – High F – High MI	86.5769	987.000	.341	.938
Verb + Adjective – Low F – Low MI	86.3462	987.008	.310	.938
Verb + Adjective – Low F – High MI	86.5385	983.732	.412	.938
Verb + Adjective – Low F – Different MI	85.6282	968.392	.511	.937
Verb + Adjective – Different F – High MI	85.8718	969.542	.469	.937

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Verb + Adverb – High F – High MI	86.1923	977.820	.528	.937
Verb + Adverb – Low F – High MI	86.4231	979.572	.488	.937
Adverb + Adjective – High F – High MI	86.5897	988.401	.308	.938
Adverb + Adjective – Low F – Low MI	86.4231	975.910	.555	.937
Adverb + Adjective – Low F – High MI	86.6667	987.758	.324	.938
Adverb + Adjective – Low F – Different MI	85.8333	961.699	.532	.936
Adverb + Adjective – Different F – High MI	86.0000	973.792	.455	.937

Table 12

Pilot Interview Findings

Patterns	Reasons/Explanation	Participants' Comments
Why certain test items and sentences were difficult:	- Sentences had many unknown words - The meaning of the sentences was not clear - Test items (the collocational matrix) included unknown words	*"Some words have several meanings, and I am familiar only with one, not suitable for that particular sentence."* *"Sometimes I choose the word because it is the only one I knew."* *"Long sentence and the end is not clear."* *"Need more context."*
Why certain test items and sentences were easy:	- Participants knew this exact phrase (have heard/read it before) - The meaning of the sentences was clear	*"To be honest, nothing was really difficult for me because I was trained to "feel" what words will be the most proper in this or that gap."* *"The only words that make sense."*
Why the test was rated at 4 ("difficult") or 5 ("very difficult")	- The instructions were not clear - The test format was new, and the task was not clear	*"In sentences where two or more word combinations were possible, I don't know what choose."* *"Know the words but cannot match them."* *"Don't know how to combine words and put in the sentences."* *"Confused. What's the differences between words?*

157

Patterns	Reasons/Explanation	Participants' Comments
Test-reading/completion strategies	- Focused on the meaning of separate words - Focused on the meaning of whole sentences - Focused on the gap and the collocational matrix - Took a guess - Compared L1 and L2	*"Read more academic words first."* *"Read first part [of the sentence]and found first word; then read second part and found the whole. Check does it make sense?"* *"In my language it's same, so I guessed right."*
Vocabulary-learning strategies:	- Learning words one-by-one - Learning words in combinations - Using authentic context - Comparing languages	*"Repetitive words from pop-culture and TV shows (mostly British)."* *"Interesting words from newspapers"* *"Expressions in my language and English"*

The interview results indicated that the test task and format were new to participants, which might have impacted test scores. As a result, more explanations and examples were provided in the guidelines for the final test version. Specifically, the sentences with a gap (retrieved from the BNC and COCA) were simplified because, as shown in the comments, participants had some difficulties with unfamiliar words and/or insufficient context. Additionally, the majority of participants rated the test a 4 ("difficult") on the Likert scale; interview comments suggest that up to 40% of the words that comprised collocations and/or their distractors were not quite familiar to the test takers. For example, participants indicated that they might have heard/read this word before but were not sure of its exact meaning and so took a guess. As it was not possible to predict which words might be unfamiliar to all participants, in order for participants not to get distracted by possible unknown items, dictionary definitions were provided for each group of synonyms in the final test version (see Appendix K).

Finally, respondents' interview responses about their test completion and vocabulary-learning strategies shed some light on factors that might be accountable for their collocational knowledge. For example, participants indicated that they focused on the meaning of words and sentences rather than on word form, and used both intuition (taking a guess) and their L1 ("In my language it's the same"). Also, they confirmed learning words both individually and in combinations in the context of media, that is, using authentic input.

Post-test survey. On the basis of the pilot study interviews, a post-test survey (see Appendix L*)* was developed. The survey contained multiple-choice and Likert scale questions designed to examine participants' background (age, first language, and how

159

long they have been studying English) and their perceptions of the test (how they would rate this test in terms of its difficulty on the Likert scale from 1 ("very easy") to 5 as ("very difficult"), and why they gave such a score).

Participants

The final test version (see Appendix K) and the post-test survey (see Appendix L) were administered to 78 participants who were advanced EAP learners from two major language backgrounds (7 Arabic speakers and 69 Chinese speakers). Two more participants chose not to identify their L1. Most of the participants were in the age range of 18-20 (n = 52) or 21-25 (n = 20), with only a few participants in the age groups of 26-30 (n = 1) and over 30 (n = 2). The average age of onset for learning English was 8.7, and the average number of years spent on learning English was 11. The majority of participants had started learning English between 5 and 10 years old, and spent more than 10 years learning it (see Figures 7 and 8).
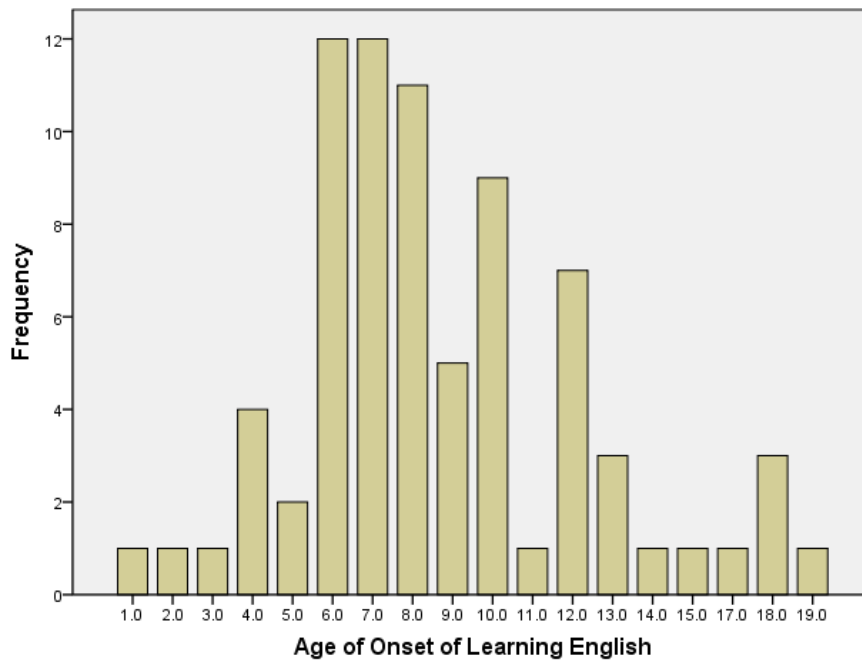
Figure 7. Age of onset of learning English.
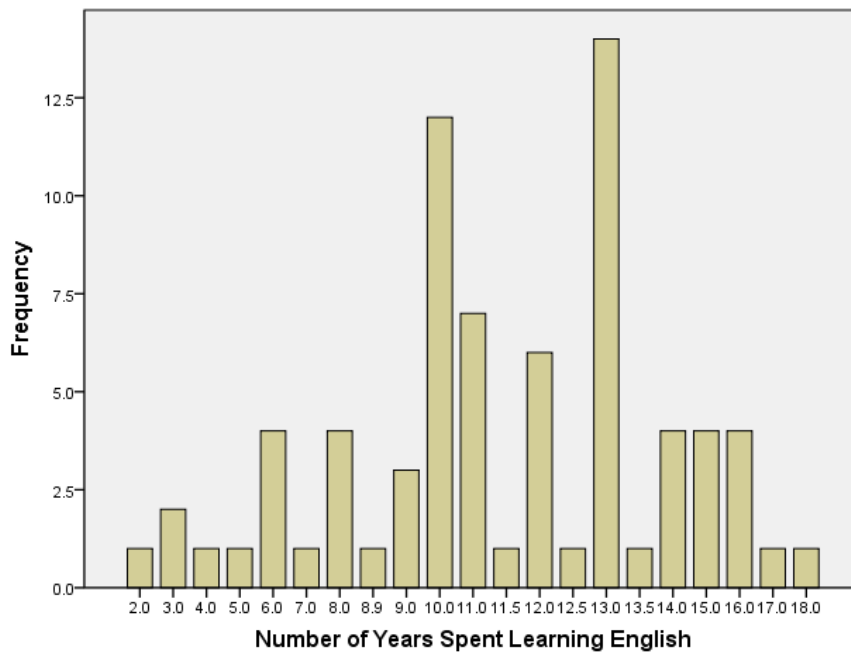
Figure 8. Number of years spent learning English.

Data Collection Procedures

The paper-based test was administered in three same-level advanced EAP classrooms. The participants were given 45 minutes to complete the test (approximately 1.5 minute per test item), based on the average time it had taken participants in the pilot test phase. Participants were free to respond to test items in a non-consecutive order and make corrections. They were specifically instructed not to dwell on one item for more than two minutes since the purpose of the test was spontaneous recognition and controlled production rather than a detailed analysis of sentences and test items. Additionally, they were forbidden to refer to any external resources such as dictionaries since that would have eliminated the purpose of the test and word definitions had already been provided for them. Although participants were not given any prior information on collocations so as to elicit spontaneous recognition and production, they were provided with explicit instructions and examples on how to approach the tasks and what to focus on (word combinations that best fit in the overall sentence context). Immediately after completing the test, participants were given 15 minutes to complete the post-test survey.

Data Analysis and Results

The data from the test and the post-test survey were coded using SPSS Version 22.0. The coding model emulated the pilot study, that is, seven main factors (low F – different MI – different MSTs, high MI – different F – different MSTs, five MSTs – different F – different MI) and their combinations were identified as independent variables (see Appendix J), and their relation to the dependent variable (overall test score) was explored through consecutive procedures of correlation and multiple regression.

Descriptive statistics. Table 13 presents the average mean scores of each of the test item categories. The average mean score on the test was 13.74 (40%) with the highest score of 26 and the lowest score of 4. The MST of Verb + Adverb had the largest percentage (47.5%) of correct responses; however, since there were only 4 items in this category, the Verb + Noun type with 9 items (44% of correct responses) could be considered as leading among the five MSTs and therefore, the most accurately recognized and produced. The Adverb + Adjective type (34.5% of correct responses) appeared the most difficult for the participants. Consequently, descriptive statistics suggests that Verb + Noun collocations might be easier for the participants than other MSTs.

Table 13

Percentage of Correct Responses for the F, MI, and MST Categories

	Adjective + Noun	Verb + Noun	Verb + Adjective	Verb + Adverb	Adverb + Adjective	Total for all MSTs
High F - High MI	1.19/3 = 39.7%	1.95/3 = 65%	0.7/2 = 35%	1.06/2 = 53%	0.66/2 = 33%	5.56/12 = 46%
Low F - Low MI	1.51/3 = 50%	1.14/3 = 38%	0.91/2 = 45.5%	0	0.84/2 = 42%	4.4/10 = 44%
Low F - High MI	0.8/3 = 26.7%	0.86/3 = 28.7%	0.71/2 = 35.5%	0.84/2 = 42%	0.57/2 = 28.5%	3.78/12 = 31.5%
TOTAL	3.5/9 = 39%	3.95/9 = 44%	2.32/6 = 38.7%	1.9/4 = 47.5%	2.07/6 = 34.5%	13.74/34 = 40%

When looking at the variations of F and MI, it is possible to deduce that collocations with low F and high MI scores (31.5% of correct responses) were the most difficult, while collocations with high F and high MI scores (46% of correct responses) were the easiest for the participants. When comparing F and MI scores across five MSTs, Verb + Noun (65% of correct responses) and Verb + Adverb (53% of correct responses) scored the highest in the high F – high MI category; Adjective + Noun (50% of correct responses) and Verb + Adjective (45.5% of correct responses) scored the highest in the low F – low MI category; and Verb + Adverb (42% of correct responses) and Verb + Adjective (35.5% of correct responses) scored the highest in the low F – high MI category. Additionally, although all five MSTs scored the lowest in the category of low F – high MI, Adjective + Noun, Verb + Adjective, and Adverb + Adjective types scored the highest in the category of low F – low MI (50, 45.5, and 42% of correct responses respectively), while Verb + Noun and Verb + Adverb collocations scored the highest in the category of high F – high MI (65 and 53% of correct responses respectively).

These findings show that high F collocations are easier to recognize and produce for ESL learners, while MI score does not appear to make a difference. This corresponds to the results of Edmonds and Gudmestad (2014), Ellis et al. (2008), and Nguyen and Webb's (2016) studies, which identified that unlike in NSs, the fluency of collocation recognition and production in ESL speakers does not seem to be impacted by MI score even though they are sensitive to the frequency of occurrence of word combinations.

It is also possible to see that recognition and production of collocations with low and high F and MI scores is distributed somewhat unevenly depending on the MSTs. Thus, less frequent and weaker Adjective + Noun, Verb + Adjective, and Adverb +

Adjective types were identified by the participants more easily than the respective Verb + Noun and Verb + Adverb Types. At the same time, it was again confirmed that strong and frequent Verb + Noun collocations were most easily identified.

The frequency counts run for the collocation-specific variables showed that kurtosis and skewness were within the norm (+/- 1). For the seven main factors, the distribution was mostly symmetrical. For the combination of factors, the majority of variables demonstrated strong negative skewness with only a few variables positively skewed (Adjective + Noun – high F – high MI; Verb + Adjective – low/high F – high MI; and Verb + Adverb – high F – high MI). This result indicates that the median is generally larger than the mode, and, consequently, there is a large distance between higher and lower test scores, which does not necessarily indicate outliers but potentially depends on participants' prior language learning background and L1.

Correlation. In order to identify a possible relationship between the test score and independent variables, a PPMCC analysis was run. The results indicated that there is a statistically insignificant small negative correlation for the test scores and the factors of the age of onset of learning English ($r = -.01$, $p = .890$) and additional languages ($r = -.18$, $p = .130$), and a statistically insignificant small positive correlation for the factors of L1 ($r = .18$, $p = .100$), age ($r = .07$, $p = .100$) and the number of years spent learning English ($r = .15$, $p = .170$). These results are due to the fact that the majority of participants shared the same L1 background (Arabic and Chinese) and age of onset, were in the same age group, and had approximately the same number of years of L2 instruction. Thus, non-collocation specific factors, which in this study would have been distractors, were eliminated.

The PPMCC for the seven main factors (F, MI, and five MSTs) identified a statistically significant strong positive correlation for the factors of low F – different MI – different MSTs ($r = .93$, p = .000), high MI – different F – different MSTs ($r = .94$, $p = .000$), and Verb + Noun – different F – different MI ($r = .79$, $p = .000$) All three combinations of F and MI scores, regardless of their MST, also received high scores: high F – high MI – different MSTs ($r = .82$, $p = .000$), low F – low MI – different MSTs ($r = .82$, $p = .000$), and low F – high MI – different MSTs ($r = .78$, $p = .000$). Only a moderate positive correlation was identified for the remaining four MSTs: Adjective + Noun ($r = .67$, $p = .000$), Verb + Adjective ($r = .62$, $p = .000$), Verb + Adverb ($r = .66$, $p = .000$), and Adverb + Adjective ($r = .61$, $p = .000$). When separated into subtypes based on their F and MI, most of the five MSTs had moderate to low PPMCC scores, ranging from .20 to .67. The only exception was Verb + Noun collocations with low F – different MI ($r = .69$, $p = .000$) and high MI – different F ($r = .74$, $p = .000$).

These results indicate that, as Durrant and Schmitt (2009), Metin and Karaoğlan (2011), and Siyanova and Schmitt (2008) suggest, the factors of F and MI are highly interrelated and hence, the potential problem of multicollinearity. Further, as identified in the descriptive statistics as well as in existing research by Boers et al. (2014), Eyckmans (2009), Koya (2005), Nesselhauf (2005), Revier (2009), and Szudarski (2016), Verb + Noun collocations could be potentially strong predictors of successful collocation recognition and controlled production and need to be focused on when teaching collocations.

Regression. For the purpose of identifying which of the seven independent variables (F, MI, and five MSTs) and/or their combinations would be the best score

predictor(s) and therefore, the most influential factor(s), a multiple regression analysis was implemented in several steps.

Step 1: Regression analysis for F, MI, and five MSTs. A standard multiple linear regression was run for the seven main factors. The model summary indicated the R-squared value of 1, which means that the model explains 100 percent of the variance in the test scores and is, therefore, valid. The value of each of the independent variables is further explained in Table 14.

Table 14

Standard Multiple Linear Regression. Independent Variables: F, MI, and MSTs

Model	Standardized Coefficients		Collinearity Statistics	
	Beta	Sig.	Tolerance	VIF
1(Constant)		1.000		
Low F - Different MI - All MSTs	.000	1.000	.092	10.820
Different F - High MI -All MSTs	.000	1.000	.067	15.029
Adjective + Noun -Different F - Different MI	.310	.000	.244	4.093
Verb + Noun - Different F - Different MI	.372	.000	.170	5.892
Verb + Adjective - Different F - Different MI	.263	.000	.329	3.038
Verb + Adverb - Different F - Different MI	.230	.000	.198	5.050
Adverb + Adjective - Different F - Different MI	.293	.000	.256	3.905

The beta values for the factors of low F – different MI – different MSTs and high MI – different F – different MSTs equaled .000 where $p = 1$, which nullified their predictive value. Additionally, as shown in the PPMCC analysis, correlations for the factors of low F – different MI – different MSTs ($r = .93$), and high MI – different F – different MSTs ($r = .94$) were approximately the same, which indicates these values are mutually exclusive (Tabachnick & Fidell, 2007). This finding shows that F and MI are interconnected rather than completely separate entities, which corresponds to the research by Durrant and Schmitt (2009), Granger and Bestgen (2014), and Siyanova and Schmitt (2008).

The beta values for standardized coefficients were more than .3 in case of two MSTs: Verb + Noun (.37) and Adjective + Noun (.31) collocations, where $p = .000$. These statistically significant findings are in accordance with the findings of the correlation analysis and indicate that Verb + Noun collocations have the largest unique contribution to the test score, though Adjective + Noun collocations could also be potentially influential. Therefore, both were retained for further analysis.

Finally large VIF values (more than 10) and small tolerance values (less than .10) for the factors of low F – different MI – different MSTs and high MI – different F – different MSTs in Table 14 suggest high inter-variable correlation and, therefore, multicollinearity. Consequently, the normal probability plot of the regression standardized residual (Figure 9) demonstrated significant deviations of the observed value from the expected value, and although the standardized residuals were roughly rectangularly and evenly distributed in the scatterplot (Figure 10), they were not concentrated in the center, which puts the predictive value of the analysis in question.

Figure 9. Normal P-P plot for independent variables: F, MI, and MSTs.
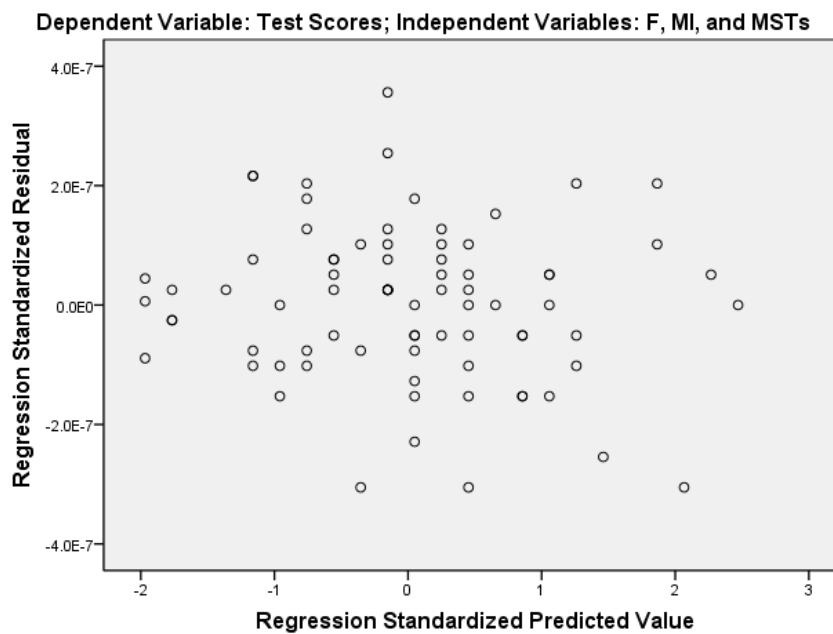
Figure 10. Scatterplot for independent variables: F, MI, and MSTs.

These findings could be related to the fact that the factors of F and MI are interconnected and as such, might predict collocation recognition and production in combination rather than separately.

In sum, at Step 1 of the regression analysis, it was identified that first, the factors of low F (different MI – different MSTs) and high MI (different F – different MSTs) are highly correlated and contribute to multicollinearity, which is in accordance with the initial PPMCC analysis discussed in the previous section. At the same time, two out of five MSTs (Verb + Noun and Adjective + Noun) appear to be strong score predictors, which was also the case in the PPMCC analysis. Finally, these preliminary findings demonstrated that the categories of F, MI, and MSTs are interconnected, hence the predictive value of F and MI was nullified when analyzed as independent variables on the same scale with the five MSTs.

However, it remained to be seen whether the categories of F and MI could be analyzed in their relation to one another as binary oppositions on the scale from low to high and independently of MSTs, and whether Verb + Noun and Adjective + Noun MSTs would remain significant if their F and MI scores were not accounted for. The next two steps of the regression analysis split seven main factors into two categories (combinations of F and MI as independent variables belonging to Category 1 and five MSTs as independent variables of Category 2).

Step 2: Regression analysis for binary oppositions low – high F and low – high MI. A standard multiple linear regression was implemented again, but this time for the purpose of identifying the predictive value of the factors of low and high F and MI alone regardless of the MSTs. The variables of F and MI were further divided into two binary

oppositions (low F versus high F and low MI versus high MI) with different PPMCCs, where Pair 1 was high F – high MI ($r = .82$) and low F – high MI ($r = .78$) and Pair 2 was low F – high MI ($r = .78$), and low F – low MI ($r = .82$). The R-squared value for Pair 1 was 88.4%, and for Pair 2, 87.6% ($p = .000$), which makes the results statistically significant. As shown in Tables 15 and 16, beta values proved significant and identical for Pair 1 and Pair 2 (.58 for the low – high F scale and .51 for the low – high MI scale) with tolerance values (more than .10) and VIF values (less than 10) within norm. The normal probability plots (Figure 11) and the scatterplots (Figure 12) had no deviations of the observed value from the expected value, and standardized residuals were evenly distributed, which indicated no data singularity or multicollinearity.

These results suggest that the binary oppositions of low – high frequency of occurrence and low – high frequency co-occurrence do play a role in collocation recognition and production when analyzed independently from one another. This finding aligns with existing research by Granger and Bestgen (2014), Siyanova-Chanturia and Spina (2015), and Szudarski and Conklin (2014), who suggest that high frequency collocations are acquired faster, and studies by Hill et al. (2000) and Nguyen and Webb (2016), who state that stronger collocations with a higher MI score might present more challenges.

Overall, the results of the regression analysis at Step 2 demonstrated that although the factors of F and MI are strongly correlated and complement one another, when analyzed independently of one another, binary oppositions of F and MI on the scale from low to high (namely, low – high F and low – high MI) can be accurately used to predict the test scores.

Table 15

Standard Multiple Linear Regression. Independent Variables: High F – High MI and Low F – High MI

Model		Standardized Coefficients		Collinearity Statistics	
		Beta	Sig.	Tolerance	VIF
1	(Constant)		.048		
	High F - High MI - All MSTs	.584	.000	.789	1.267
	Low F - High MI - All MSTs	.517	.000	.789	1.267

Table 16

Standard Multiple Linear Regression. Independent Variables: Low F – High MI and Low F – Low MI

Model		Standardized Coefficients		Collinearity Statistics	
		Beta	Sig.	Tolerance	VIF
1	(Constant)		.000		
	Low F - High MI - All MSTs	.510	.000	.775	1.290
	Low F - Low MI - All MSTs	.580	.000	.775	1.290

Figure 11. Normal P-P plots for Pair 1 (top: High F – High MI and Low F – High MI) and Pair 2 (bottom: Low F – Low MI and Low F – High MI).

Dependent Variable: Test Scores; Independent Variables: High F - High MI and Low F - High MI
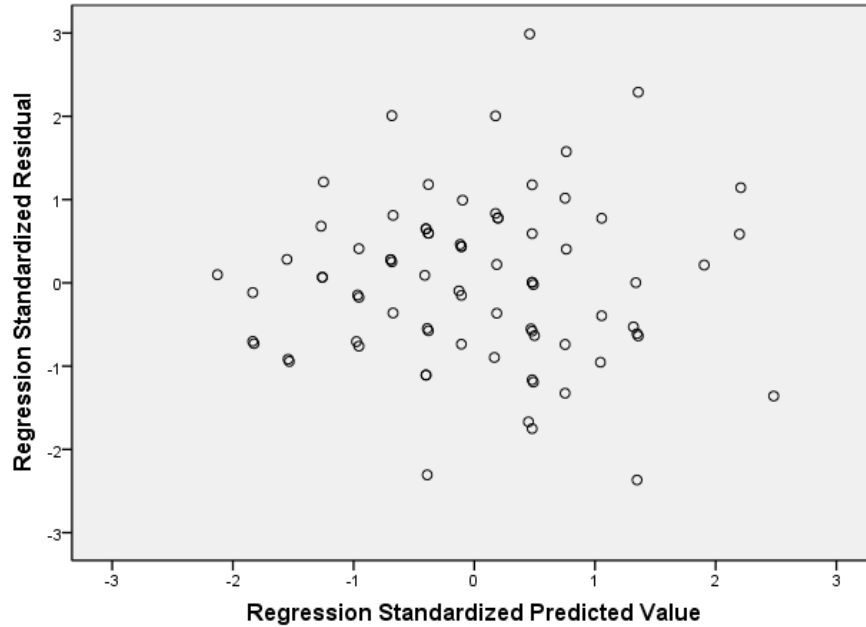
Dependent Variable: Test Scores; Independent Variables: Low F – Low MI and Low F – High MI
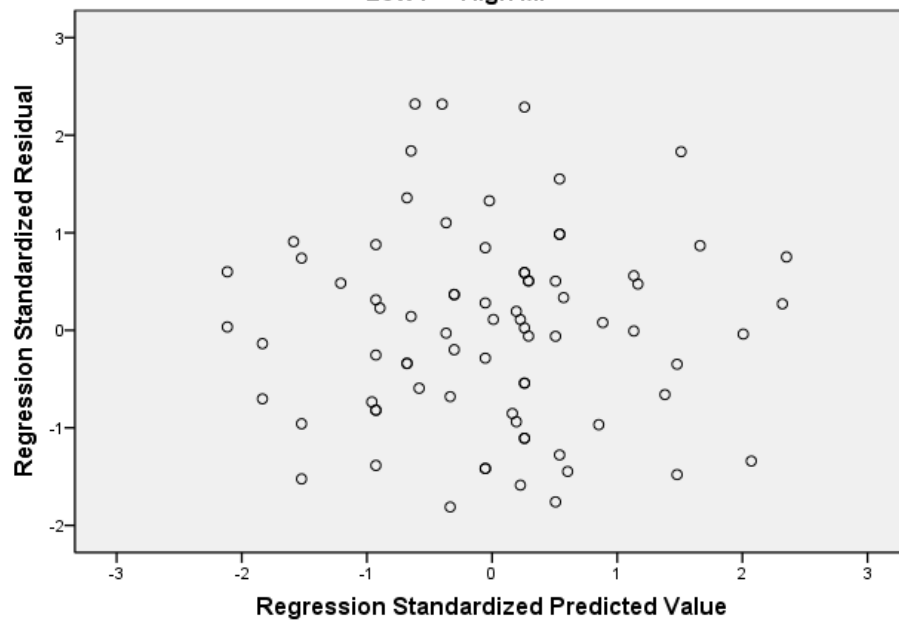
Figure 12. Scatterplots for Pair 1 (top; High F – High MI and Low F – High MI) and Pair 2 (bottom; Low F – Low MI and Low F – High MI).

Step 3: Regression Analysis for five MSTs. A standard multiple linear regression was run for five MSTs regardless of their F and MI in order to confirm that Verb + Noun and Adjective + Noun collocations are indeed the strongest score predictors. The model summary indicated an R-squared value of 1, which means that the model explains 100 percent of the variance in the test scores and is, therefore, valid. The value of each independent variable is further explained in Table 17.

Table 17

Standard Multiple Linear Regression. Independent Variables: Five MSTs

Model		Standardized Coefficients		Collinearity Statistics	
		Beta	Sig.	Tolerance	VIF
1	(Constant)		1.000		
	Adjective + Noun	.310	.000	.804	1.243
	Verb + Noun	.372	.000	.609	1.641
	Verb + Adjective	.263	.000	.752	1.330
	Verb + Adverb	.230	.000	.622	1.609
	Adverb + Adjective	.293	.000	.804	1.243

The beta values for standardized coefficients were more than .3 in the cases of two MSTs: Adjective + Noun (.31) and Verb + Noun (.37) with p = .000, which corresponds to the results of the standard multiple linear regression analysis run with seven factors and indicates that these two MSTs have the largest unique contribution to the test score. The VIF values (less than 10) and adequate tolerance values (more than .10) for the five MSTs suggested no multicollinearity. However, although the

standardized residuals were evenly and rectangularly distributed in the scatterplot (Figure 14), the observed value still exhibited large differences from the expected value in the normal probability plot (Figure 13), which somewhat decreases the predictive value of the analysis. In other words, Adjective + Noun and Verb + Noun collocations appear to be strong score predictors only to some extent and further analysis is needed to confirm their role.

Normal P-P Plot of Regression Standardized Residual

Dependent Variable: Test Scores; Independent Variables: Five MSTs

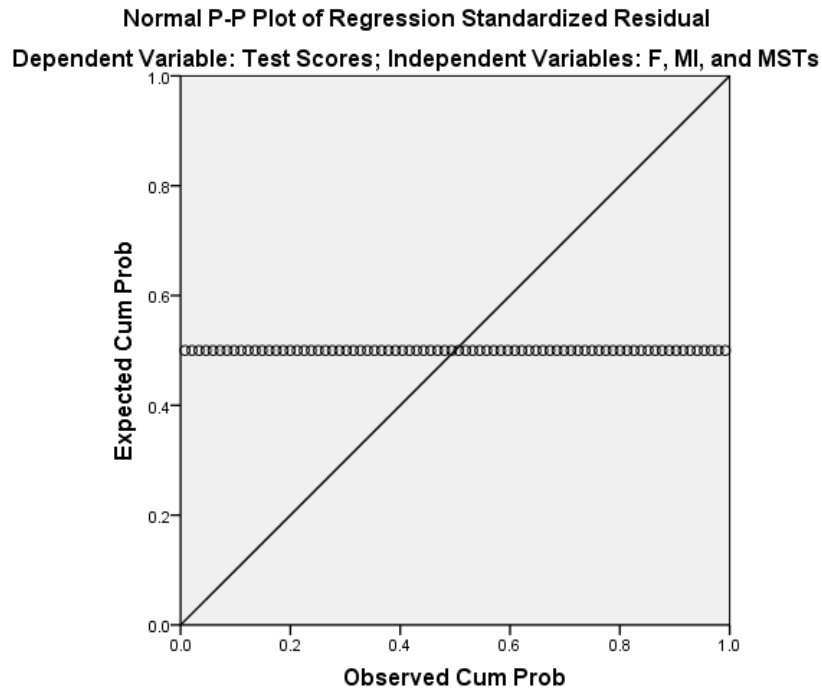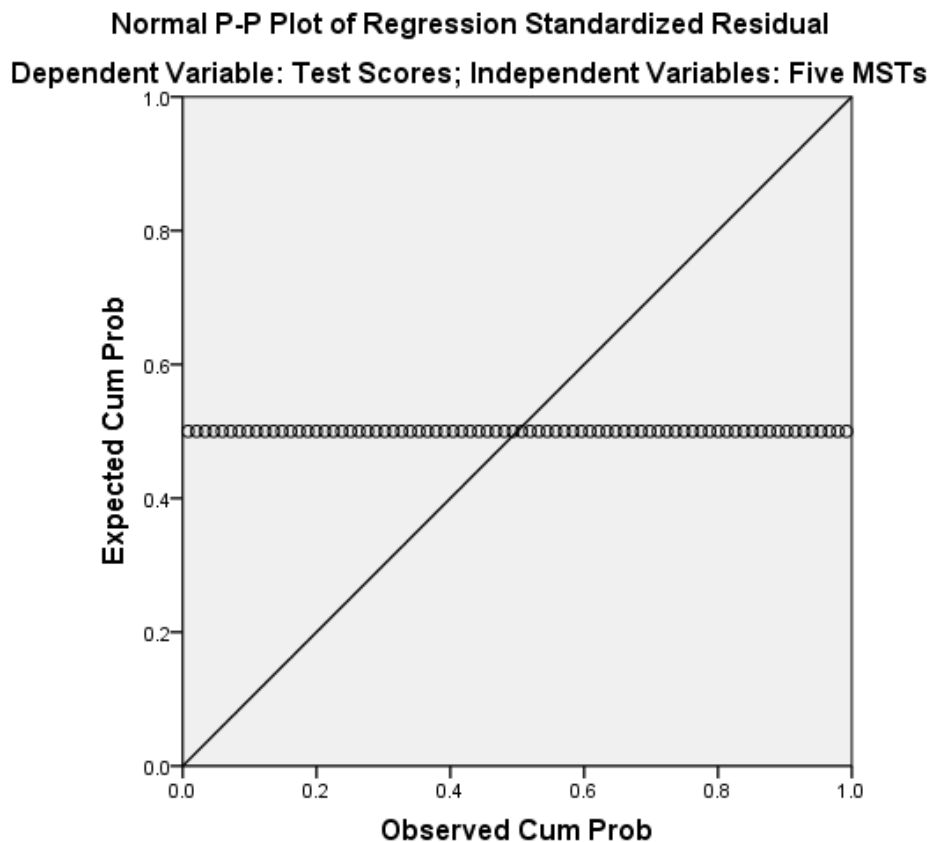*Figure 13.* Normal P-P plot for independent variables: five MSTs.
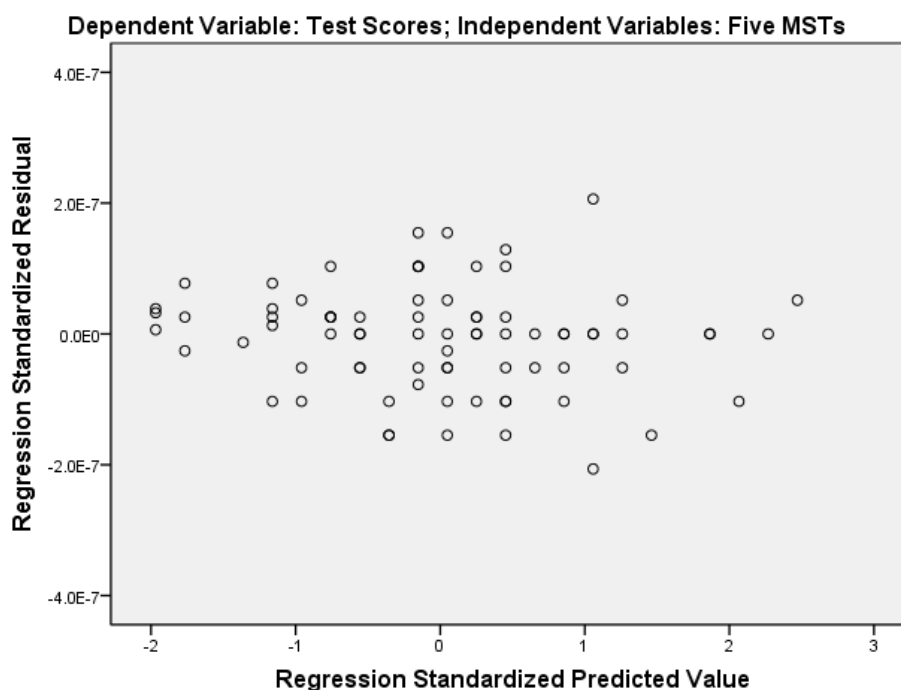
Figure 14. Scatterplot for independent variables: five MSTs.

In order to confirm these somewhat contradictory initial findings, a stepwise regression was run for the five MSTs. Similarly to the results of the standard multiple linear regression, Verb + Noun ($r = .79$) and Adjective + Noun ($r = .67$) collocations were identified as the strongest score predictors, while Adverb + Adjective ($r = .60$) and Verb + Adjective ($r = .62$) types were identified as the weakest predictors. At the same time, adjusted R-squared values demonstrated that the consecutive addition of Verb + Noun collocations to Adjective + Noun collocations and then to Verb + Adjective collocations made the largest difference to the percentage of predictability (from 63% to 79%, and to 90% respectively). In other words, Adjective + Noun collocations increased test validity by 16 percent, and Verb + Adjective collocations by 11 more percent. It

appears that even though Verb + Adjective collocations were not strongly correlated to test scores ($r = .62$) and had a low beta value (.26), they still might make a difference.

These results suggest that the three MSTs (Verb + Noun, Adjective + Noun, and Verb + Adjective) emerge as significant score predictors. While Verb + Noun and Adjective + Noun collocations have been identified as important in multiple research studies by Boers et al. (2014), Jaén (2007), Pellicer-Sánchez (2017), Revier (2009), Siyanova and Schmitt (2008), and Szudarski (2016), the MST of Verb + Adjective requires more analysis and explanation.

Finally, a standard multiple linear regression was implemented for these three factors only in order to further clarify their significance and predictive value. The model summary indicated an R-squared value of .95, which means that the model explains 95 percent of the variance in the test scores and is, therefore, valid. The value of each of the independent variables was further explained in Table 18.

The beta values predictably increased in comparison to the previous models (.55 for Verb + Noun; .37 for Adjective + Noun; and .34 for Verb + Adjective) while the tolerance values (more than .10) and VIF values (less than 10) were normal, indicating no multicollinearity. This finding is further confirmed by the normal probability plot in which the observed value is aligned with the expected value (Figure 15), and the scatterplot (Figure 16) in which the standardized residuals are rectangularly distributed and somewhat evenly concentrated along the same horizontal 0 line. These results confirm that the three MSTs of Verb + Noun, Adjective + Noun, and Verb + Adjective collocations are significant score predictors; consequently, it can be concluded that collocation acquisition depends on the MST.

Table 18

Standard Multiple Linear Regression. Independent Variables: Adjective + Noun, Verb +

Noun, and Verb + Adjective

Model		Standardized Coefficients		Collinearity Statistics	
		Beta	Sig.	Tolerance	VIF
1	(Constant)		.229		
	Adjective + Noun	.370	.000	.833	1.201
	Verb + Noun	.555	.000	.820	1.220
	Verb + Adjective	.348	.000	.870	1.149

Normal P-P Plot of Regression Standardized Residual

Dependent Variable: Test Scores; Independent Variables: Verb + Noun, Adjective + Noun, and Verb + Adjective
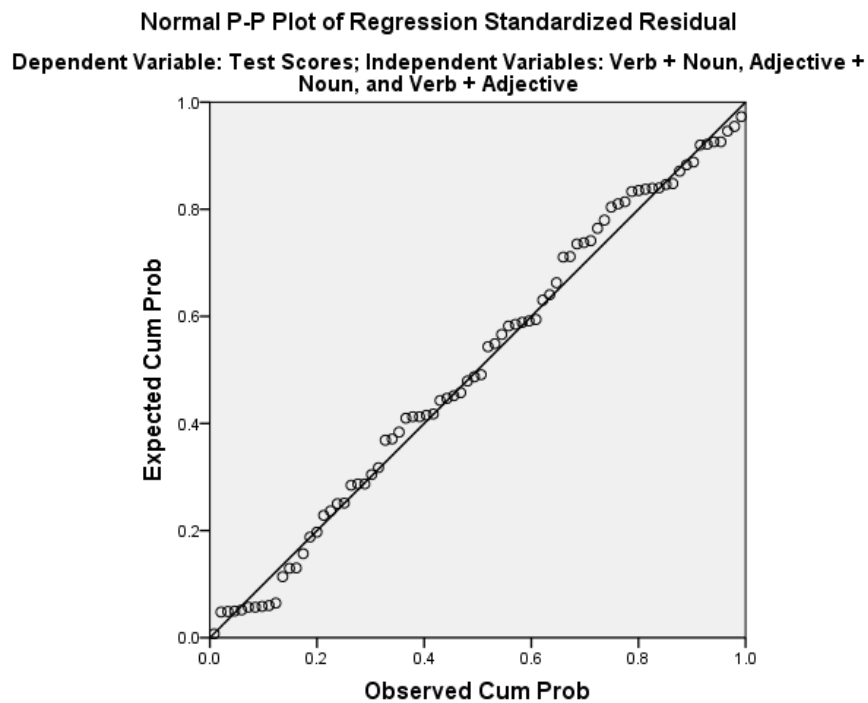
Figure 15. Normal P-P plot for independent variables: Adjective + Noun, Verb + Noun,
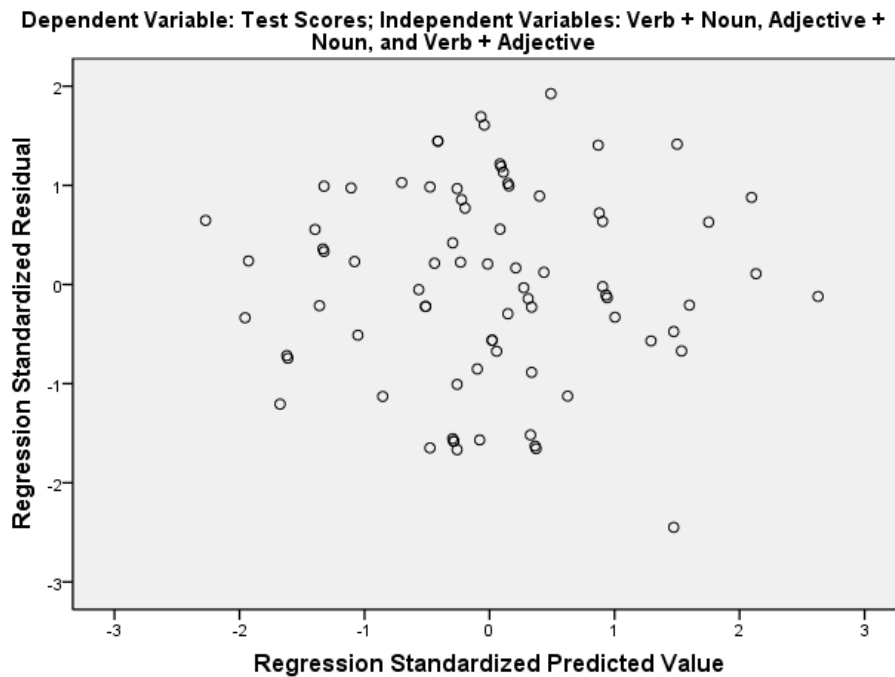
and Verb + Adjective.

Figure 16. Scatterplot for independent variables: Adjective + Noun, Verb + Noun, and Verb + Adjective.

Step 4: Regression analysis for the three MSTs and their differences in terms of low – high F and low – high MI scores. While the factors of F and MI when taken as independent variables on the same scale with five MSTs had a null predictive value (because of their high correlation and consequently, multicollinearity of the model), their binary oppositions of low – high F and low – high MI still emerged as significant when analyzed separately (see Step 2). Therefore, it was deemed worthwhile to examine whether three MSTs (Verb + Noun, Adjective + Noun, and Verb + Adjective) might still be strong score predictors or predict the score differently if we control for the possible effect of their low – high F and/or low – high MI scores.

In order to address this question, a hierarchical multiple regression was implemented with the three MSTs as the main factors and the binary oppositions of low – high F and low – high MI as the controlled factors, where:

1. <u>High F</u> – high MI and <u>low F</u> – high MI, where the binary opposition of low – high F was controlled factor 1; and

2. Low F – <u>low MI</u>; and low F – <u>high MI</u>, where the opposition of low – high MI was controlled factor 2.

The results of the analysis of controlled factor 1 (Table 19) indicated that the factor of low – high F accounted for 88.4% of the variance in the test scores; where after controlling for it, the total model accounted for 95.7 % of the test score variance and predictability [$F(5,72) = 323$, $p = .000$]. In other words, the three MSTs explained an additional 7.3% of the variance in test scores, where the R-squared value changed by .07 (from .88 to .95), [F change $(3,72) = 41$, p = .000]. Only in the case of Verb + Noun collocations did the statistically significant beta value (beta = .32, p = .000) exceed the minimum threshold of .3. Additionally, in the normal probability plot, the observed value is in line with the expected value (Figure 17), which increases the predictive value of the analysis. The standardized residuals (Figure 18) were rectangularly distributed and evenly concentrated along the same horizontal 0 line in the scatterplot (Tabachnick & Fidell, 2007). Both plots suggested linearity and normality of the analysis model.

These results demonstrate that although the factor of frequency of occurrence might be significant; the MST might play a bigger role, particularly in the case of Verb + Noun collocations, thus making this MST the strongest score predictor, which corresponds with the previous stages of analysis.

Table 19

Hierarchical Multiple Regression. Independent Variables: Low – High F and Three

MSTs (Adjective + Noun, Verb + Noun, and Verb + Adjective)

Model		Standardized Coefficients Beta	t	Sig.
1	(Constant)		2.011	.048
	High F - High MI - All MSTs	.584	13.209	.000
	Low F - High MI - All MSTs	.517	11.696	.000
2	(Constant)		.046	.963
	High F - High MI - All MSTs	.278	7.116	.000
	Low F - High MI - All MSTs	.251	6.893	.000
	Adjective + Noun - Different F - Different MI	.220	6.986	.000
	Verb + Noun - Different F - Different MI	.328	8.894	.000
	Verb + Adjective - Different F - Different MI	.195	6.240	.000
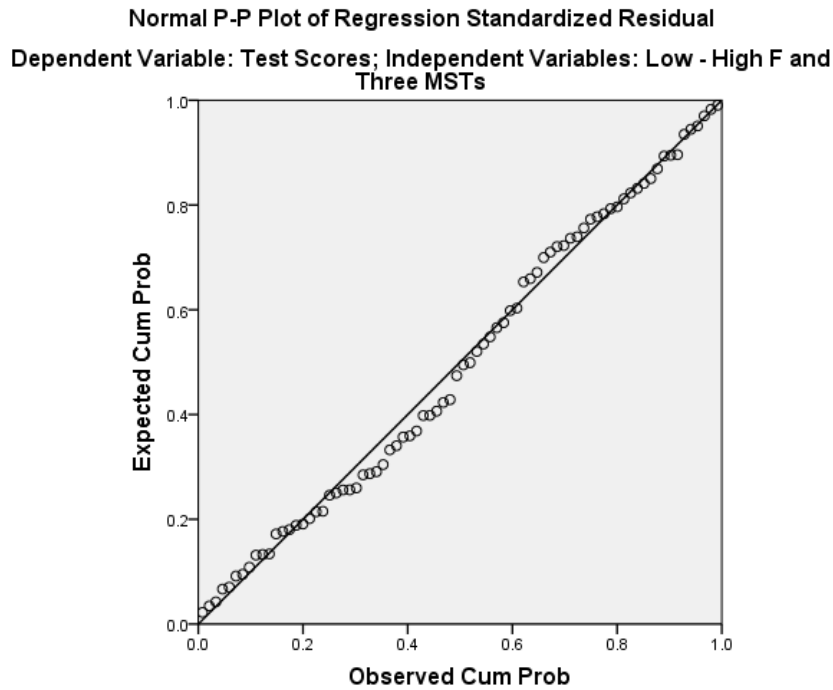
Figure 17. Normal P-P plot for Low – High F and three MSTs (Adjective + Noun, Verb + Noun, and Verb + Adjective).
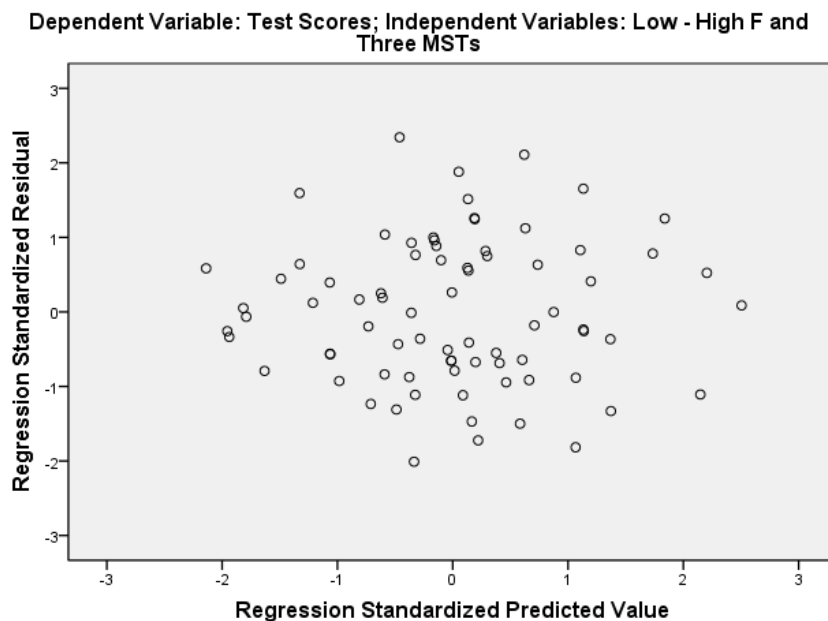
Figure 18. Scatterplot for Low – High F and three MSTs (Adjective + Noun, Verb + Noun, and Verb + Adjective).

Similarly to the analysis of low – high F, the results of the analysis of controlled factor 2 (Table 20) demonstrated that 87.6% of the variance in test scores could be explained by the controlled factors of low – high MI, and after entering the factors of the three MSTs, the total model explained 94.1% of the test score variance and predictability [$F(5,72) = 227$, $p = .000$]. That means that the control variables (the three MSTs) explained an additional 6.4% of the variance in test scores, where the R-squared value changed by .06 (from .88 to .94), [F change $(3,72) = 25$, $p = .000$]. Only the statistically significant beta value of Verb + Noun collocations (beta = .33, $p = .000$) exceeded the minimum threshold of .3. Similarly to the analysis of controlled factor 1 (the binary opposition of low – high F in the three MSTs), in the normal probability plot for controlled factor 2 (the binary opposition of low - high MI in the three MSTs) the observed value was in line with the expected value (Figure 19), which signifies the good predictive value of the analysis, while the standardized residuals for controlled factor 2 (Figure 20) were rectangularly distributed and spread out along the horizontal 0 line in the scatterplot, which signifies linearity and normality of the analysis model.

These results suggest that although similarly to the factor of frequency of occurrence, low and high frequency of co-occurrence might have some influence on collocation acquisition, the MST of Verb + Noun collocations is more important in terms of statistical significance.

Table 20

Hierarchical Multiple Regression. Independent Variables: Low – High MI and Three

MSTs (Adjective + Noun, Verb + Noun, and Verb + Adjective)

Model		Beta	t	Sig.
		Standardized Coefficients		
1	(Constant)		4.701	.000
	Low F - High MI - All MSTs	.510	11.051	.000
	Low F - Low MI - All MSTs	.580	12.566	.000
2	(Constant)		1.826	.072
	Low F - High MI - All MSTs	.250	5.654	.000
	Low F - Low MI - All MSTs	.213	4.006	.000
	Adjective + Noun - Different F - Different MI	.233	5.865	.000
	Verb + Noun - Different F - Different MI	.335	6.768	.000
	Verb + Adjective - Different F - Different MI	.232	6.309	.000

Normal P-P Plot of Regression Standardized Residual

Dependent Variable: Test Scores; Independent Variables: Low - High MI and Three MSTs
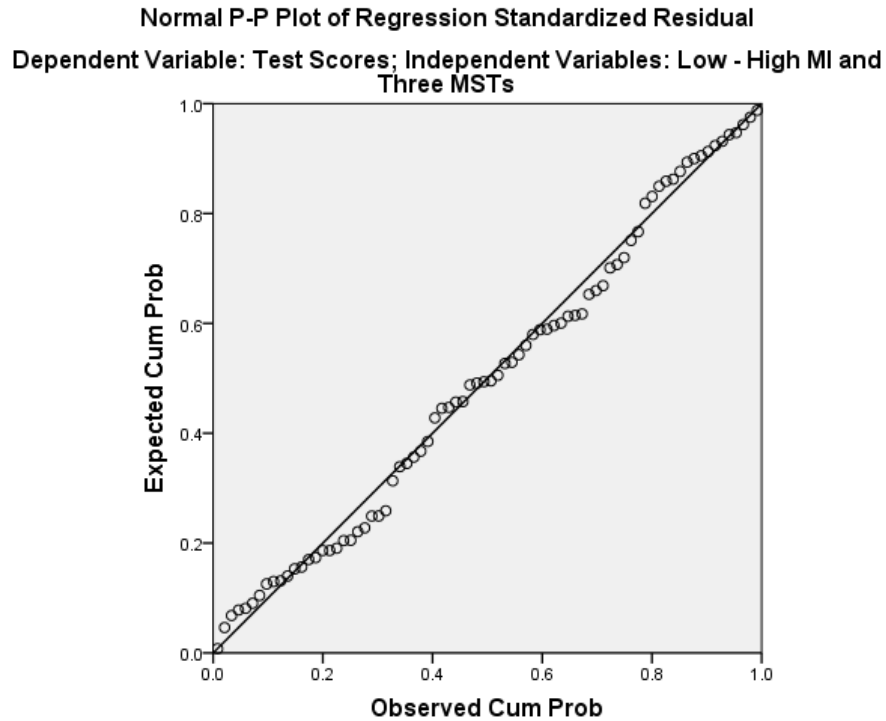
Figure 19. Normal P-P plot for Low - High MI and three MSTs (Adjective + Noun, Verb + Noun, and Verb + Adjective).

Dependent Variable: Test Scores; Independent Variables: Low - High MI and Three MSTs
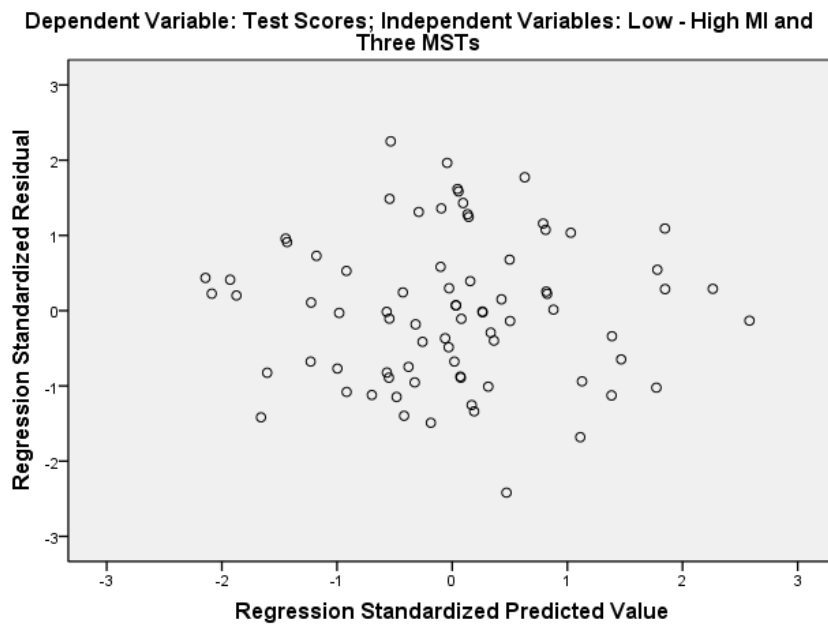
Figure 20. Scatterplot for Low – High MI and three MSTs (Adjective + Noun, Verb + Noun, and Verb + Adjective).

The overall results of the hierarchical regression analysis with (1) low – high F and (2) low – high MI as controlled factors and three MSTs as control measures showed that Verb + Noun collocations with both low – high F and low – high MI remained significant score predictors. That means that compared to this MST, the frequency of occurrence and co-occurrence might not be the determining factors in terms of the score predictability as was initially identified in Step 1 of the regression analysis.

Post-test survey. Since the post-test survey collected information on participants' age and language background, the variables of age, first language, age of onset of learning English, and number of years spent on learning English were included as an extra set of independent external (non-collocation-specific) variables that, according to the research (Foster, 2009; Granena & Long, 2013; Groom, 2009; Liao, 2010; Schmitt, 2000; Wray, 2002), might be potentially significant. However, the descriptive statistics for continuous variables eliminated these factors as having unequal distribution and violating the assumption of normality because the majority of participants belonged to the same age and L1 group, started learning English at approximately the same time, and spent an equal number of years learning English at school ($SD = .64$; 1.37; 3.69; and 3.33 respectively, $p = .000$).

Another post-test survey question focused on participants' previous vocabulary learning strategies and asked them about their preferred way of learning words. Most respondents (slightly over 60%) answered that they were familiar with and were used to learning words in combinations/patterns, which might indicate that they have encountered and memorized formulaic multiword units.

Participants' responses on the last question of the survey, "Please rate this test on the scale from 1 ("very easy") to 5 ("very difficult") and explain why you gave such a score", were analyzed separately from the test itself. A total of 40 participants (51.3%) rated the test at 3 ("moderately difficult"), 33 participants (42.3%) rated the test at 4 ("difficult"), and 3 participants (3.8%) rated it at 5 ("very difficult"). The accompanying comments mostly resembled the results of the semi-structured interview at the pilot stage of this study in that they related to an unfamiliar task format (e.g., "Not familiar with this kind of practice") and difficulty when choosing the right words (e.g., "I know the meaning but I don't know which one is better"), which might be connected to limited collocational competence, i.e., knowledge of what words connect together, as identified by Hill (2000), Pellicer-Sánchez (2017), and Szudarski and Carter (2016).

Discussion

The results of Stage II of the research indicate that even advanced English language learners with on average over 11 years of experience learning English in different environments have relatively low collocational competence, where the average test score was only 13.74 (40%). These results are not surprising since many researchers (Groom, 2009; Han, 2004; Nesselhauf, 2005; Nguyen & Webb, 2016; Wray, 2008) indicate that the development of collocational knowledge is slow for second language learners; in the case of adult language learners, the process of fossilization can take place, or else analytical processing of word combinations can slow down their learning. However, fossilization and analytical processing might not be the only mitigating factors. The descriptive statistics identified a negative skewness for most variables, which presupposes a large difference between high and low test scores in many categories.

Indeed, the highest test scores were above 20, while the lowest test scores were as low as 4. This might be an indicator of students with different language proficiency being in the same academic level in EAP programs, an issue that is beyond the scope of this study.

The purpose of this research stage was to determine if such characteristics of collocations as low – high frequency of occurrence, low – high frequency of co-occurrence, or five MSTs (Adjective + Noun; Adverb + Adjective; Verb + Adjective; Verb + Adverb; and Verb + Noun) could be test score predictors and as such, influence collocation acquisition, namely active recognition and controlled production. For this purpose, the items of a collocation-targeted test partially modelled after Revier (2009) were selected because they belonged to these five MSTs and had a certain score on the scale of low to high F and MI. Thus, it was possible to single out and analyze each of the above-mentioned factors (and their different combinations) as independent variables in their relation to overall test scores.

The initial findings of the descriptive statistics on the test scores were similar to previous studies (Durrant & Schmitt, 2010; Evert, 2008; Gitsaki, 1996; Koya, 2005; Poulsen, 2005; Sonbul, 2015; de Souza Hodne, 2009; Szudarski & Conklin, 2014; Wolter & Gyllstad, 2013) in that they confirmed an important role that frequency of occurrence plays in the active recognition and controlled production of collocations. The highest number of correct responses (46%) fell into the category of high F collocations, while the lower number of correct responses (37%) belonged to the category of low F collocations. Consequently, we can conclude that frequent collocations are more easily recognized and produced while less frequent collocations present more challenges for ESL learners.

Nevertheless, as other studies (Ellis et al., 2008; Nguyen & Webb, 2016; Revier, 2009) have suggested, the frequency of co-occurrence measured by the strength of association cannot be written off as an insignificant factor or be equated to the frequency of occurrence, even though these two factors might be interconnected. As explained in studies by Ackermann and Chen (2013), Ellis et al. (2008), and Fellbaum (2007), it is the MI score rather than the raw/normed F per million score that serves to differentiate between non-fixed word combinations and collocations because MI demonstrates that word components of a lexical unit do not co-occur by chance. A high MI score indicates that when combined together, these words have a certain distinctive meaning and, as Firth (1957) long ago confirmed, are considered as a whole inseparable meaningful unit, that is, a collocation. This statement was confirmed when selecting test items, conducting the pilot study, and refining the test format. The word combinations with higher MI scores can have low F as in *disposable income* (F = 1.62, MI = 12.66) or high F as in *ethnic minority* (F = 10.1, MI = 10.07) and still be easily identified as collocations, but the lower the MI score, the less fixed and less obvious a collocation becomes. Thus, despite approximately the same F scores, the collocation *give access* (F = 1.03, MI = 3.76) is less fixed and consequently, allows for more variation than the collocation *cast doubt* (F = 1.75, MI = 10.07). Additionally, it appears that there is a negative correlation between MI and F scores, where the higher the F score in a word combination with a low MI score, the less collocation characteristics are exhibited in this particular word combination. Thus, the word combination *new information* (F = 10.37, MI = 3.56) allows for variation and cannot be considered as fixed.

The descriptive statistics demonstrated that collocations with higher MI scores, regardless of their low or high frequency, account only for 39% of correct responses, while collocations with lower MI scores account for 44% of correct responses; therefore, collocations with higher MI scores are less easily recognized and produced and present more challenges to the learners. These findings correspond to the results of Nguyen and Webb's (2016) study that also identified a negative correlation between MI scores and learners' collocation recognition, and to some extent agree with the findings of Ellis et al.'s (2008) research, which determined that unlike native speakers, learners of English could not process formulaic units with higher MI scores. It appears, then, that ESL learners cannot recognize and produce strong collocations very well precisely because of their high level of formulaicity.

As Firth (1957) and Ellis et al. (2008) mention, highly formulaic units possess a distinctive and somewhat metaphorical meaning, allow for less variability, and have to be perceived as a whole, which might be a problem in the case of adult ESL learners in particular because of their analytical rather than holistic processing of word combinations (Wray, 2002). For example, the collocation *disposable income* has a high MI score of 12.66, where the word *disposable* is not used in its literal meaning of "no longer useful; one-time use only"; in connection to its node *income*, it acquires a new figurative meaning as "readily available". One can compare this to the collocation *present evidence*, which barely crosses the minimum MI threshold of 3; in this example, both words in this collocation, as well as the whole word combination, are used in their literal meaning and thus variability is possible, for example, *current evidence* or *existing evidence*.

The descriptive statistics also identified interesting results for the five MSTs. As most researchers (Boers et al., 2014; Eyckmans, 2009; Gyllstad, 2009; Koya, 2005; Peters, 2016) have suggested, Verb + Noun collocations are difficult to acquire and therefore, are most frequently addressed in the research. However, in this study, test scores were the highest for Verb + Noun collocations (44%), where within this category, high F – high MI Verb + Noun collocations (such as *provide evidence* or *collect data*) accounted for 65% of the total score, which suggests that this MST might actually be the easiest to recognize and produce. This finding is important because it contradicts some studies but aligns with research by Wolter and Yamashita (2015), who determined that Verb + Noun collocations are processed faster than Adjective + Noun collocations; likewise, this research found that Verb + Noun collocations are not particularly problematic for ESL learners. However, the factor of frequency of occurrence might be partially responsible for this result in the study, which suggests interrelatedness of the factors of frequency and morphosyntactic structure.

Other MSTs that are not typically the focus of research, such as Verb + Adjective (38.7% of correct responses) or Adverb + Adjective (34.5% of correct responses), accounted for lower percentages of correct responses and therefore could be more problematic for learners. This finding is difficult to explain, because of the general lack of research on these MSTs, and because there is only a small and somewhat contradictory difference in the average MI and F scores among these three MSTs, with Verb + Noun collocations on the one hand, and Verb + Adjective and Adverb + Adjective collocations on the other hand. For example, the average frequency is 4.21 for Verb + Noun and 4.44 for Verb + Adjective collocations, yet the average MI is 7.02 for Verb + Noun and 6.37

for Verb + Adjective collocations, which makes it appear as if a higher MI score might be responsible for the differences in test scores for these two categories. This finding is in accordance with Nguyen and Webb's (2016) study about correlation between MI and test scores. However, the MI score was the largest in case of Adverb + Adjective collocations (MI = 7.45); despite this, this MST had the lowest total score of 34.5%.

The correlation and multiple regression analysis clarified this puzzling finding to some extent. External factors (age, number of years participants have been learning English, and age of onset of learning English) predictably did not emerge as statistically significant and did not correlate with test scores due to a general lack of data variability (the majority of the participants belonged to the same L1 background and age group, and, consequently, spent approximately the same amount of time learning English). These factors have already been observed and discussed at Stage I of the research, and would have been distractors at the present stage, which focuses on collocation-specific factors only.

Correlation analysis showed a statistically significant high positive correlation (over the cutoff of .7), which is considered high enough by Pallant (2010) for such independent variables as low F – different MI – different MSTs, high MI – different F – different MSTs, and all three combinations of F and MI (high F – high MI, low F – low MI, and low F – high MI). Moreover, since the PPMCC was approximately the same for the combinations of F and MI, this might suggest that these variables are highly interconnected. The results of the correlation analysis were confirmed by the standard multiple linear regression analysis of the seven main factors of F, MI, and five MSTs. Due to their interconnectedness, the variables of F and MI skewed the picture. Although

F and MI could impact the test score and lead to the R-squared change (from 88% to 96%), their beta coefficient values were less than .3, while Sig. values equaled 1. Additionally, high VIF values (more than 10) and low tolerance values (less than .10) for the seven main factors contributed to multicollinearity. Therefore, when analyzed as independent variables on the same scale with five MSTs, F and MI nullified one another and did not have any predictive value as individual separate variables.

This finding corresponds to previous research by Barnbrook et al. (2013), Metin and Karaoğlan (2011), and Siyanova and Schmitt (2008), who argue that the frequency of occurrence and co-occurrence (measured by the strength of association) are essentially different techniques for measuring the same thing. However, this finding contradicts studies (Durrant & Schmitt, 2010; Fellbaum, 2007; Ngyuen & Webb, 2016; Wolter & Gyllstad, 2013) that suggest that the frequency of occurrence and co-occurrence signify different characteristics of collocations and, as such, measure different things (namely, that the frequency of occurrence measures how often this word combination can be found in a corpus, and the frequency of co-occurrence confirms that this word combination has not co-occurred by chance and has a certain distinctive meaning).

Nevertheless, the subsequent standard multiple linear regression analysis demonstrated that low and high F and low and high MI could be strong test score predictors when considered as binary oppositions. The first binary opposition of <u>high F</u> – high MI and <u>low F</u> – high MI had respective PPMCCs of .82 and .78, and beta values of .58 and .51 with no multicollinearity diagnosed. These results confirm the findings that emerged earlier in the descriptive statistical analysis and those explored in studies by Durrant and Schmitt (2010), Siyanova-Chanturia & Spina (2015), Sonbul (2015),

Szudarski and Conklin (2014), and Wolter and Gyllstad (2013). Notably, there is a statistically significant difference in ESL learners' perception of more and less frequent collocations. Moreover, ESL learners acquire frequent collocations more effectively than non-frequent collocations and, consequently, the former are less challenging for ESL learners because of the amount of previous exposure. Based on these two points, it is possible to say that low – high frequency of occurrence might serve as a score predictor and, therefore, impact collocation recognition and controlled production.

The second binary opposition was that of high MI – low F and low MI – low F collocations that had PPMCCs of .78 and .82, and beta values of .51 and .58 with no multicollinearity diagnosed. These results indicate that although both high and low MI scores might be score predictors, a high MI score is a slightly less significant predictor. This might not be surprising in light of the descriptive statistics analysis that, similar to Nguyen and Webb's (2016) research, identified a negative correlation between MI score and test scores. However, this finding contradicts Ellis et al.'s (2008) study, which claimed that there was no significant difference in the way ESL learners process low and high MI-scored formulaic units. As shown in the present research, there is clearly a difference in high and low MI-scored collocation recognition and production because of different beta values and correlation coefficients; furthermore, this difference is statistically significant, with both variables being score predictors.

It is important to note that the two pairs of binary oppositions have identical PPMCCs and beta values. This can be partially explained by the fact that the two pairs share a common variable (low F – high MI with r = .78 and beta value = .51). Even so, it remains somewhat puzzling why the other two different variables (low F – low MI and

high F – high MI) would emerge as having an identical PPMCC of .82 and beta value of .58. A possible explanation is that in these combinations there are F and MI bands that are approximately similar (both are either low or high) and thus they even each other out, which again is a sign of their interconnectedness (Metin & Karaoğlan, 2011; Siyanova & Schmitt, 2008).

As demonstrated, the results for the F and MI scores are somewhat inconclusive because even though F and MI scores seem to accurately predict test scores when separated into high and low binary oppositions, they are still highly interrelated and a certain claim cannot be made about their individual role in collocation recognition and controlled production, nor whether they should be considered as separate entities or combined as one.

The picture is clearer and more promising in the case of the five MSTs. The correlation analysis confirmed the results of the descriptive statistics, where Verb + Noun collocations were highly correlated with test scores ($r = .79$, $p = .000$) and the remaining four types scored below the threshold of .70. These findings were further explained by the standard multiple linear regression analysis that identified Verb + Noun collocations as the strongest score predictors with a statistically significant beta value of .37. This finding is important because although a vast majority of the researchers (Boers et al., 2014; Eyckmans, 2009; Gyllstad, 2009; Koya, 2005; Revier, 2009) have focused on this MST when teaching, testing, or analyzing collocations, only Nguyen & Webb (2016) and Peters (2016) have attempted to compare it with another MST (Adjective + Noun collocations). While Nguyen and Webb (2016) do not come to conclusive findings that

might indicate that these two MSTs might be test score predictors, Peters (2016) suggests than MST might indeed impact collocation recognition and recall.

The regression analysis showed that Adjective + Noun collocations also accounted for the test scores. Although their PPMCC was only .67, their beta value was .31, which makes this MST a statistically significant score predictor. This finding again does not agree with Nguyen and Webb's (2016) comparative study of MSTs; however, it does agree with other studies on Adjective + Noun collocations (Jaén, 2007; Pellicer-Sánchez, 2017) in the sense that it also identifies this MST as potentially problematic for learners of English. In addition, the descriptive statistics indicated that the category of Adjective + Noun collocations had only 39% of correct responses, which is five percent less than the percentage of correct responses in the category of Verb + Noun collocations.

A further important finding was that Verb + Adjective collocations also emerged as potential score predictors as a result of the stepwise multiple regression. Although this MST had only 38.7% of correct responses (equal to the percentage of responses in the Adjective + Noun category), its PPMCC was only .62 and its beta value was .26, where the adjusted R-squared value demonstrated that it increased test validity by 11 percent, and thus contributed to the model. Verb + Adjective collocations have not been not widely explored in the research with the exception of Almela's (2011) study, which described the characteristics of this particular MST and identified that these collocations are the result of a compound, i.e., an Adjective + Noun collocation embedded in Verb + Noun collocation (p. 43). In this sense, the significance of Verb + Adjective collocations, in spite of their relatively low PPMCC and beta value, makes sense. This particular MST is a modified version of two other MSTs (Verb + Noun and Adjective + Noun), and as

197

such shares many of their characteristics and might complement them as a score predictor. For example, such collocations as *make explicit* represents a potential combination of Verb (*make*) + Noun (*goals, intentions, statements,* etc.) and Adjective (*explicit*) + Noun (*goals, intentions, statements,* etc.).

The subsequent standard multiple linear regression run for the factors of Verb + Noun, Adjective + Noun, and Verb + Adjective collocations confirmed these findings. All three factors had statistically significant increased beta values of .55, .37, and .34 respectively, and the new model explained 95 percent of the variance in test scores, which signifies its validity. This finding is of great importance since to the best of my knowledge, previous research studies did not compare MSTs in terms of their role in collocation acquisition; therefore, the present research is among the first to identify these three MSTs as influential in terms of collocation recognition and controlled production.

Since the binary oppositions of low – high frequency and low – high MI scores were identified as significant score predictors when analyzed independently of MSTs, an additional research question emerged: Would the above-mentioned MSTs still play a role when predicting test scores after normed frequency per million and mutual information were controlled for? The hierarchical multiple regression was run for Verb + Noun, Adjective + Noun, and Verb + Adjective collocations as control measures in two steps:

1. With the binary opposition of low – high normed frequency per million as two controlled factors. For example, Adverb + Adjective collocations *particularly appropriate* (F = 1.57) versus *slightly different* (F = 9.6); and

2. With the binary opposition of low – high mutual information score as two controlled factors. For example, Adverb + Adjective collocations *relatively common* (MI = 4.25) versus *mutually exclusive* (MI = 13.24).

The results of the standard multiple linear regression align with what was identified at Step 1 of the analysis. It was confirmed that binary oppositions of low – high frequency of occurrence and frequency of co-occurrence measured by the MI score impact test scores to some extent when analyzed in the context of MSTs because on the one hand, after controlling for different F or MI, two MSTs (Adjective + Noun and Verb + Adjective) lose their predictive value altogether. On the other hand, regardless of their low – high F or low – high MI, Verb + Noun collocations still emerged as significant score predictors with beta values above the minimum threshold of .30. This means that the MST might be a more significant score predictor than the frequency of occurrence and co-occurrence.

Overall, the results of the data analysis at Stage II of the research suggest that when taken as independent variables and compared to MSTs, neither of the factors (low – high frequency of occurrence or low – high frequency of co-occurrence) seems to have a decisive role in active recognition and controlled production of collocations because they are highly interrelated. These findings appear to confirm the point of view of researchers who considered frequency of occurrence and co-occurrence as parts of a whole –that is, different statistical techniques measuring the same thing (Barnbrook et al., 2013; Metin & Karaoğlan, 2011; Siyanova & Schmitt, 2008).

At the same time, when analyzed as binary oppositions, regardless of the MSTs, low – high normed frequency per million and low – high mutual information scores do

appear to have some significance. Additionally, we found out that high frequency collocations account for higher test scores, while low frequency collocations seem to correspond to lower test scores. This finding agrees with research done by Ellis et al. (2008), Siyanova-Chanturia & Spina (2015), Sonbul (2015), Szudarski and Conklin, (2014), and Wolter and Gyllstad (2013), who found that ESL learners are susceptible to the frequency of occurrence of formulaic units and consequently, recognize and produce high frequency items more easily. Quite the opposite seems to be the case with the frequency of co-occurrence; while collocations with higher MI scores account for the lower test scores, collocations with lower MI scores are correlated with the higher test scores. A possible explanation and confirmation for this phenomenon can be found in Nguyen and Webb's (2016) study, namely that stronger collocations are recognized and produced less easily than weaker collocations by ESL learners due to their lack of variability.

While findings on frequency and MI are interesting and worth mentioning, this study's most significant and innovative contribution to collocation research is its comparison of the five MSTs. Previously, the majority of researchers focused on analyzing one particular MST and argued for either Verb + Noun (Boers et al., 2014; Gyllstad, 2009; Revier, 2009) or Adjective + Noun (Jaén, 2007; Pellicer-Sánchez, 2017) collocations being the cornerstone of collocational knowledge. Further, those who described the morphosyntactic classification of collocations did not attempt any comparison in terms of the predictive value of different MSTs. For example, Nguyen and Webb (2016) compared Verb + Noun and Adjective + Noun collocations but did not identify them as strong score predictors. In the present study, Verb + Noun collocations

were identified as the strongest score predictors, which at the same time are the most easily recognized and produced since they accounted for the highest percentage of correct responses on the test. Adjective + Noun collocations emerged as a close runner-up in terms of their ability to predict test scores, but at the same time, they appear to be more problematic for ESL learners. This finding contradicts the results of Peters' (2016) research that concluded that adjective forms are less variable than verb forms and hence present less of a challenge for learners, and yet aligns with Wolter and Yamashita's (2015) study that suggested that Verb + Noun collocations are processed faster than Adjective + Noun collocations. Wolter and Yamashita (2015) explained this finding by the fact that participants shared the same L1 and collocational items were selected based on the principle of congruency, which might have impeded the results. Since the predominant majority of participants in the present study share the same L1 (Chinese), the factor of congruency might also play a role and needs to be further explored in future studies. Another runner-up is the MST of Verb + Adjective collocations, which are not frequently addressed in collocation research. However, this finding can be explained in the context of Almela's (2011) study, which suggests that Verb + Adjective collocations are compounds of Verb + Noun and Noun + Adjective collocations.

Summary and Implications for Stage III

Chapter 6 described Stage II of the research, which focused on internal collocation-specific factors, such as the frequency of occurrence measured by the normed frequency per million, the frequency of co-occurrence (that is, the strength of association measured by the mutual information (MI) score), and five morphosyntactic types (Adjective + Noun, Adverb + Adjective, Verb + Adjective, Verb + Adverb, and Verb +

Noun) as potentially influential in collocation acquisition. The factors of frequency of occurrence and co-occurrence emerged as significant in previous studies, while the factors of five MSTs had not previously been explored in their relation to one another. The additional factors of age, primary languages, age of onset of learning English, and number of years spent learning English were not accounted for because this was not the focus of this stage of the research and because of the lack of diversity in participants' background data. However, since some of these (for example, length of residence in an English-speaking country or L1 congruency) had emerged as significant either at Stage I of this research or in other studies, they might have had a potential minor impact on the test scores.

An academic collocation recognition and controlled production-targeted test modelled after Revier (2009) was administered to 78 participants, advanced EAP learners from a predominantly Chinese L1 background and in the age group of 18-25. The test used the format of Collocation Constituent Matrix grid (Revier, 2009); test items were selected from the Academic Collocation List (Ackermann & Chen, 2013) on the basis of their normed frequency, MI score, and morphosyntactic structure; test sentences were retrieved from the BNC and COCA; and synonyms for nodes and collocates served as distractors.

The results confirmed to some extent the predictive value of the binary oppositions of low – high frequency of occurrence and low – high frequency of co-occurrence (MI score). Additionally, the study identified that a high frequency of occurrence accounted for higher test scores, which means high frequency collocations are easier to recognize and produce in comparison to low frequency collocations. However, a

low MI score contributed to higher test scores compared to a high MI score, which suggests that stronger collocations more difficult for ESL learners to acquire. These findings correspond to other studies that focused on frequency and mutual information (Edmonds & Gudmestad, 2014; Nguyen & Webb, 2016). Another important finding was that three out of five MSTs (Verb + Noun; Adjective + Noun; and Verb + Adjective) appear to be strong score predictors with Verb + Noun collocations being the most significant. However, highly frequent and strong Verb + Noun collocations also appeared to be the least problematic for ESL learners which contradicts some previous studies (Boers et al., 2014; Gyllstad, 2009; Revier, 2009).

On the basis of this research stage, a few questions emerged, namely (1) whether teaching collocations based on the factors of frequency of occurrence, frequency of co-occurrence, and five MSTs and then testing learners' recognition and controlled production might lead to similar or different results; and (2) whether such teaching might be effective and feasible. Until answers to these two questions are confirmed, it cannot be claimed with certainty that the above-mentioned collocation characteristics are indeed influential factors. Chapter 7 will describe and provide a rationale for the collocation-targeted teaching module tasks and activities that were performed at Stage III of the research, and Chapter 8 will focus on the pre- and post-teaching collocation recognition and controlled production-based test administered at Stage III.

Chapter 7: Stage III. Collocation-targeted Teaching Module

Rationale and Design

Although collocation-specific factors were examined as potentially influential in collocation acquisition in Chapter 6, it remains to be seen whether teaching collocations based on these factors is feasible and will lead to any significant differences in collocation acquisition. For this purpose, a set of teaching activities and learning tasks and exercises based on three collocation-specific factors (frequency of occurrence, frequency of co-occurrence, and five MSTs) was developed and administered in beginner-level EAP classrooms (44 participants in total). Teaching took place over two weeks, twice per week, 30-40 minutes per class. This chapter will focus on providing the rationale for the exploratory collocation-based teaching module and describing the lesson structure, tasks, and exercises.

Several important decisions were made while developing teaching activities and exercises. First, I compared the criteria of collocation selection implemented in other studies and identified what comprises collocational knowledge and consequently, collocational errors. Second, it was necessary to choose collocational items to teach. Finally, an effective teaching approach had to be chosen. These decisions are discussed below.

Criteria of Collocation Selection

Most researchers focus on the frequency of occurrence (F) of either whole collocations (Boers et al., 2014) or the node word (Nizonkiza & Van de Poel, 2014; Pellicer-Sánchez, 2017). Those who opt for node frequency usually select words from a list of the most frequent (BNC or COCA-based) words. Some researchers also consider

the frequency of co-occurrence measured by mutual information (MI), i.e., strength of association (Ellis et al., 2008; Szudarski & Carter, 2016) and what it stands for, namely semantic and morphosyntactic links between collocational components (Nation & Newton, 1997). The next important criteria is the morphosyntactic type (MST); however, most studies stop at either Verb + Noun (Boers et al., 2014; Eyckmans, 2009; Gyllstad, 2009; Koya, 2005; Revier, 2009;) or Adjective + Noun (Jaén, 2007; Pellicer-Sánchez, 2017), or both (Nguyen & Webb, 2016; Peters, 2016; Wolter & Yamashita, 2015) since these two types are traditionally considered the most problematic for L2 learners. Only a few researchers take on more than two MSTs (Gitsaki, 1996; Mahvelati & Mukundan, 2012; de Souza Hodne, 2009; Szudarski & Carter, 2016). Some studies focus on grammatical collocations (Abedi & Mobaraki, 2014), while others choose lexical collocations exclusively (Attar & Allami, 2013). Additional criteria might include genre and register (neutral/general English or academic English) and L1/L2 congruency (Carrol et al., 2016; Nesselhauf, 2003; Yamashita & Jiang, 2010).

When selecting collocation items for teaching, the present study focused on the same 34 academic English collocations retrieved from the Academic Collocation List (ACL) by Ackerman and Chen (2013) that were used for testing purposes at Stage II of the research. This list incorporated the dimensions of three collocation-specific factors: the frequency of occurrence measured by the normed frequency per million, the frequency of co-occurrence or the strength of association measured by the MI score, and morphosyntactic types. Thus, collocations belonged to five MSTs: Adjective + Noun, Adverb + Adjective, Verb + Adjective, Verb + Adverb, and Verb + Noun, and had a different score on the scale of F and MI (from low to high). All collocations were short

span, that is, consecutive two-lexical-word items (Vechtomova et al., 2003), in order to avoid the mitigating factor of collocational span. A complete list of collocations can be found in Appendix M.

The rationale behind the decision to test and teach the same collocation set at both stages of the research was the necessity to (1) further explain the somewhat inconclusive findings of Stage II, namely if the factors that emerged as strong and weak score predictors will continue to be so after instruction; (2) examine how/if collocation-specific factors-targeted instruction affects learners' collocation recognition and controlled production patterns; and (3) identify whether certain collocations might be more or less easily taught depending on their semantic and structural characteristics.

Collocational Knowledge and Types of Collocational Errors

When designing collocation-targeted teaching activities, it was important to identify patterns of collocational errors so as to be able to effectively prevent them. Schmitt (2013) identified "amount of use", "speed of use", and "accuracy/appropriacy of use" errors. The "amount of use" errors are usually related to learners' overuse of certain collocations (Eyckmans, 2009; Jiang, 2009) or, conversely, their underuse (Nesselhauf, 2003). The "speed of use" errors are related to overall fluency, language proficiency level, and consequently, intuition (Kuo, 2009). The "accuracy/appropriacy of use" errors are the most frequent and can be further subdivided into several categories. The first category deals with learners not knowing the combinatory rules and morphosyntactic restrictions between collocational components. As a result, learners might use the wrong parts of speech, function words, and/or word forms, such as tense and singular/plural verb/noun forms (Nation & Newton, 1997; Nesselhauf, 2003; Wang and Good, 2007).

The second category concerns semantic restrictions. Not knowing (or not remembering) the exact words that comprise a collocation, learners resort to lexical strategies of approximation, such as substituting a word with its synonym or a descriptive/explanatory phrase translated word-for-word from their L1, or even blending collocations (Boers et al., 2014; Miyakoshi, 2009; Smith, 2005).

Since some researchers (Miyakoshi, 2009; Schmitt, 2013) identified that the largest percentage of collocational errors are "accuracy of use" erroneous patterns, the present research mostly focused on this category of errors. A focus on this type of error would allow to further explore the lexical characteristics of collocation, such as strength of association and MSTs. Moreover, "amount of use" and "speed of use" errors are based on external factors, such as intuition, L1 congruency/interference, and language proficiency level; as such, these errors cannot be reduced in a short time frame and are beyond the scope of this stage of the research.

Following Stage II of the research, within the "accuracy of use" category, only lexical errors, such as blending or paraphrasing collocations, were considered of particular interest, while grammatical errors, such as an incorrect use of prepositions, articles, singular/plural noun and verb forms, were mostly ignored. This decision was made because grammatical error patterns belong to a different domain than the three factors of frequency of occurrence and co-occurrence and MSTs, and some researchers (Nesselhauf, 2005; Wanner et al., 2006; Wanner et al., 2017) even debate on whether to consider them part of collocational competence.

Teaching Approach

As a natural continuation of the two main theoretical approaches to collocational studies –that is, distributional and phraseological approaches as discussed in Chapter 2– there have long been opposing collocation-targeted teaching approaches. While some researchers (Brandl, 2008; Graves, 2006) claim that collocations do not have distinctive features and as such, are generally non-teachable and best learned by doing (e.g., memorizing, repeating/practicing, and imitating L1 speakers' language use), other researchers (Lewis, 2000; McCarthy & O'Dell, 2005; Woolard, 2005) believe that collocations have a number of characteristics, and consequently, can be taught by noticing, analyzing, and developing awareness of these characteristics. In various studies, the former approach is called "incidental" (Schmitt, 2000), "direct" (Nation & Newton, 1997), or "input-flood" (Mahvelati & Mukundan, 2012) teaching, and is mostly supported by distributionalists (corpus-based researchers, see Chapter 2). The latter approach is favored by phraseologists (descriptive study-based researchers, see Chapter 2) and referred to as "explicit" (Nattinger & DeCarrico, 1992; Nesselhauf 2005), "lexical" (Lewis, 2000), or "input-enhancement" (Szudarski & Carter, 2016) teaching.

An incidental approach to teaching collocations is mostly intuition-based since the ability to automatically recognize and produce collocations is considered the most significant aspect of acquisition. Researchers and educators who promote this approach do not focus on collocation characteristics, nor do they address the relations between collocation components; instead, they perceive collocations as a functional pragmatic unit that needs to be retained in memory and intuitively reproduced in different genres and communicative situations. For example, the task guidelines in Marks and Wooder's

(2007) textbook state: "Choose the most natural-sounding words" (p. 38). However, learners are not given any specific criteria to help them differentiate between natural and unnatural-sounding word combinations.

In this method, learning tasks are based on the principles of language immersion (Schmitt, 2000), frequent repetition (Alali & Schmitt, 2012; Durrant & Schmitt, 2010; Pellicer-Sánchez, 2017), and inferencing/guessing/building associations from the context (Nation & Newton, 1997; Schmitt, 2000) while reading and/or speaking. At the same time, supporters of the incidental approach seem to disagree about potential usefulness of such tasks. As Schmitt (2000) states, it is not always possible to ensure language immersion and the necessary amount of repeated exposure to native-like collocations. Additionally, there is no agreement as to whether repetition is effective, and how much repetition is required. Durrant and Schmitt (2010) found that repeated exposure positively impacts collocation learning in terms of fluency. Similarly, Webb et al. (2013) stated that repetition of 15+ times positively affected collocation reception and production. On the contrary, Pellicer-Sánchez's (2017) study identified no difference between 3 and 8-times frequency exposure to collocations through reading, which means that more frequent exposure does not necessarily lead to any improvement in collocation recognition and production.

The reason for the questionable impact of repetition might have been indicated in Nation and Newton's (1997) research. They noticed that although guessing and associative networks that are built through repetition and exposure play a role in the collocation-recognition process, the mechanism of building mental associations is different for L1 and L2 speakers. While for L1 speakers it is intuition-based, L2 learners

establish links between words and/or infer from contextual clues through meticulous analysis, which means that the process of guessing is not quite incidental for them. Similarly, Edmonds and Gudmestad (2014) and Ellis et al. (2008) confirm that L2 learners are less sensitive to collocations' strength of association than L1 speakers. Therefore, Ellis (1995) explains that L2 learners benefit from explicit "deep processing"-targeted vocabulary learning strategies and tasks, which would be based on repetition and evoke certain mental images that enhance learners' ability to build associative networks. Schmitt (2010) adds that L2 learners need to be not only frequently exposed to but also actively engaged with the words and word combinations through explicit tasks that involve input enhancement and focus the learners' attention on the form, meaning, and functions of collocations.

Drawing on these findings, the explicit approach is largely analysis-based, where researchers and educators draw learners' attention to word combinations' form (Abedi & Mobaraki, 2014; Nesselhauf, 2003), meaning (Attar & Allami, 2013; Lewis, 2000), or both (McCarthy & O'Dell, 2005; Lackman, 2011; Woolard, 2005). The proponents of the explicit approach argue that since L2 learning is mostly an analytical process (Forsberg & Fant, 2010), learners often know the meaning and form of individual words (Kuo, 2009; Schmitt, 2000), but cannot connect them and therefore need to develop their collocation awareness and ability to notice links between words. Although contextual clues and associations are important, the focus of the explicit method is on the analysis of collocations as whole units (Boers et al., 2014) and their parts/components (Nesselhauf, 2003).

During the last decade, studies have attempted to compare explicit and implicit approaches and often opt for a blended approach. Mahvelati and Mukundan (2012), Sonbul and Schmitt (2013), Szudarski and Carter (2016), and Szudarski and Conklin (2014) compared the implicit approach and input-enhanced/explicit approach to teaching collocations while reading. They identified that the explicit approach works better in terms of helping learners develop their noticing ability (they are explicitly asked to pay attention and demonstrated what it is they should pay attention to), but the combination of implicit and explicit strategies might be the most effective. Lackman (2011) similarly concluded that both approaches serve their purpose. According to his study, once learners progress to a more advanced stage, they will leave the classroom and acquire and expand their vocabulary in a natural way by interacting with native speakers (that is, they will learn language incidentally); however, to reach that point, they will have had to develop a certain basis of noticing-cognitive processing skills, which explicit collocation-analysis-based instruction can help them with.

Since a blended approach is considered the most effective and least time-consuming, the present study's teaching module implemented both explicit and implicit strategies based on the following principles:

1. *Moving from analytical to holistic perception of collocations* (Barfield, 2009; Forsberg & Fant, 2010; Wray, 2002). Although the majority of adult language learners need to analyze the links between words in order to understand how collocations work and tend to split word combinations, the end goal is for them to be able to recognize and produce whole units correctly and fluently – that is, process collocations holistically. Researchers (Lai, 2005; Ying &

O'Neill, 2009) propose moving from the initial noticing/discovery stage through the analytical cognitive and consolidation stages to memorization and automatic recognition/production.

2. *Noticing* (Jiang, 2009; Lewis, 2000; Nobahar, 2017; Ying & O'Neill, 2009). Participants' attention is drawn to collocations and their characteristics through positive input enhancement, such as underlining and analyzing collocations or their parts, and negative input enhancement, such as identifying incorrect collocations. According to Ahranjani and Shadi (2012), these two types of input and noticing contribute to the learners' understanding of collocation features in the L2, and are particularly effective when dealing with "accuracy of use" errors. Ying and O'Neill (2009) added that noticing tasks also help participants acquire strategies of effective collocation learning (for example, revising non-native-like collocations in their writing, compiling personalized collocation dictionaries, or comparing collocations in L1 and L2).

3. *Inferencing from context* (Nation & Newton, 1997; Schmitt, 2000). Several tasks in the teaching module required participants to guess the node based on its collocates or vice versa, while other activities focused on guessing a correct collocation based on the surrounding lexical elements in the context of a sentence. Such tasks are particularly effective for increasing learners' knowledge and understanding of the links between words that comprise collocations.

4. *Explaining and analyzing, or "deep processing"* (Ellis, 1995; Hill et al., 2000; Woolard, 2000). Although researchers are unanimous when arguing for exploration rather than explanation when learning and teaching collocations, they also confirm that learners need some basic information on collocation features. Despite this consensus, there is a mismatch between theory and practice in most collocation-focused textbooks (Marks and Wooder, 2007; Lackman, 2011; McCarthy & O'Dell, 2005). In the first few chapters, they usually cover collocation functions, morphosyntactic classifications, and even strength of association; however, learners are rarely asked in practice exercises to analyze and reflect on these characteristics. The present study tried to avoid this common pitfall by building learning activities upon collocation characteristics, prompting participants to analyze each collocation item, and providing them with explicit explanations and feedback.

5. *Repetition* (Alali & Schmitt, 2012; Pellicer-Sánchez, 2017; Wang & Good, 2007). Since researchers hold different views on how many repetitions (from 3 to 15+ times) are required to be effective for learning (Szudarski & Conklin, 2014; Webb et al., 2013), on average, participants in this study were exposed to each of the 34 collocations at least 4 times (once per lesson) in different tasks in order to encourage retention of the item and thus, improve learners' collocational competence.

Materials

Reading-comprehension and practice exercise texts were retrieved and adapted from the COCA and BNC corpora. The majority of sentences were taken from academic

articles in such disciplines as communication studies, psychology, or journalism, and focused on general topics connected to student life, without any discipline-specific terminology. The decision to use corpora rather than readily available texts from collocation textbooks (Lackman, 2011; Lewis, 2000; Marks and Wooder, 2007; McCarthy & O'Dell, 2005; Schmitt & Schmitt, 2005; Woolard, 2005) was motivated by the necessity to provide authentic L1 input and context.

The other type of source for learning tasks and information on collocation characteristics were online collocation dictionaries, concordances, and databases such as Just-The-Word (JTW), Pro-Writing Aid: Free Online Collocation Dictionary, LexTutor, and WordAndPhrase. Using these tools, participants were able to perform such tasks as finding collocates for a node or vice versa and analyzing collocations in terms of their frequency, MST, and strength of association. Although some researchers (Akinci & Yildiz, 2017; Koo, 2006) suggest using actual academic corpora such as the BNC and COCA in the classroom, this study's teaching module opted for online resources that are simple to use and have highly interactive interface features that do not require participants to have any specific background in linguistics.

Rationale and Description of Teaching Tasks and Activities

The four lesson plans and materials built around the above-mentioned five principles are summarized in Appendix N. They focused on 34 academic collocations (see Appendix M), where the meaning of individual words comprising these collocations was expected to be familiar to participants prior to instruction. Even though learning definitions of individual words contradicts to the principle of holistic thinking (Forsberg & Fant, 2010; Wray, 2002), it was necessary in order to eliminate the factor of unknown

vocabulary. Although the principles of communicative and context/text-based teaching

(Ho & Henry, 2014; Nunan, 1991; Richards & Rogers, 2014) suggest including larger

chunks of text for reading/listening-comprehension, limited instruction time and the

necessity to focus on specific collocational characteristics (the frequency of occurrence

and co-occurrence and MSTs) motivated the decision to include sentence-level rather

than paragraph-level textual tasks, such as matching, providing acceptability judgement,

and building associative networks.

Lesson 1. The pre-instruction and post-instruction collocation-targeted test and

the post-test survey showed that although some participants were used to learning words

in combinations, the majority of participants were still relatively new to the concept of

collocation. Therefore, Lesson 1 focused on introducing and explaining collocations and

their communicative and socio-pragmatic functions. It was important not to confuse

participants by using complex terminology and to present the concept of collocation in a

way that learners would see how significant and pervasive collocations are in all forms

and genres of oral and written communication (Gitsaki, 1999; Handl, 2009; Schmitt,

2010), particularly in academic English. Additionally, Lesson 1 was to ensure learners

understood how they can actually use collocations to make their speaking and writing

more fluent, concise, and accurate (Hill, 2000; Schmitt, 2010; Siyanova & Schmitt, 2008;

Wang & Good, 2007). For this purpose, participants were engaged into the following

tasks:

Parallel comparison (adapted from Hill et al., 2000). Participants were provided

with correct and incorrect examples of collocations and asked to determine which word

combinations sounded natural to them, and why. On the basis of this discussion, the

instructor provided the basic explanation of collocation as a combination of words that always (or often) go together.

Associative networks (adapted from Nation and Newton, 1997; Revier, 2009; and Schmitt, 2000). Participants were given a list of frequent polysemantic verbs (nodes), such as *give, collect*, *keep,* and *take*, and asked to come up with a range of words that could combine with these verbs. The results led to a discussion about differences between (a) free word combinations (for example, *take a spoon/a pen/a bag*) that allow for variations, and fixed word combinations (for example, *take a break* (= rest), *take the lead* (= accept the responsibility or start winning), or *take the hint* (= understand an indirect message) that are less semantically transparent than the first group of words and more tied together; (b) lexical collocations (for example, *break the law*) and other fixed multiword units (for example, phrasal verbs such as *break up*), where the first type of formulaic unit consists of two lexical items and an article, and the second type contains one lexical item (verb) and one grammatical item (preposition); and (c) more and less frequent collocations.

Noticing and finding synonyms (adapted from Jiang, 2009; Jafarpour et al., 2013; and Sinclair et al., 2010). Participants were provided with sentences retrieved from the BNC and COCA. Each of these sentences contained one or more collocations. The task was first, to underline word combinations that participants considered fixed and useful, and second, to change the word combinations using synonyms or descriptive phrases while retaining the meaning of the original sentence. The subsequent discussion led to further consolidation of the concept of collocation, understanding of the importance of links between words, and clarification of collocation functions.

Matching (adapted from Lewis, 2000; and McCarthy & O'Dell, 2005). Participants were given a set of words that represented nodes and their collocates belonging to five MSTs and ranging in F and MI scores. Participants were asked to combine these words into two (or three)-word-long collocations, which served to reinforce their awareness of links between collocation components.

Lesson 2. Lesson 2 built on participants' knowledge of what collocations are and focused on different MSTs. The rationale behind introducing the five MSTs is that such types as Verb + Noun and to some extent Adjective + Noun and Verb + Adjective collocations emerged as the strongest score predictors at Stage II of this research. Additionally, although other researchers (e.g., Boers et al., 2014; Koya, 2005; Pellicer-Sánchez, 2017) identified Verb + Noun and Adjective + Noun collocations as the most problematic for learners, Stage II of this research indicated that these two types were actually more easily recognized by participants and therefore more attention should be paid to the other three MSTs. Lesson 2 also drew learners' attention to the structure of collocations, namely a node and dependent collocate (Sinclair et al., 2004). The following activities were included in Lesson 2:

Noticing (modelled after Hill et al., 2000; and Schmitt and Schmitt, 2005). Participants were asked to re-read the set of sentences containing collocations from the previous lesson and determine what parts of speech these collocations consist of and what the main and supporting words are. This exercise led to a discussion and more examples of five basic MSTs and the concept of a node and collocate(s).

Finding collocates for a node (modelled after Schmitt and Schmitt, 2005; and Woolard, 2005). Participants were given a list of node words and asked to determine

what parts of speech these words represent. Then, using online collocation databases, they were asked to find collocates for each node. In order to ensure their understanding of the morphosyntactic links between the words, participants were encouraged to search for collocates that belonged to particular parts of speech. For example, they were given the noun *majority* and asked to look for adjectives that might combine with this noun.

Concordance lines (adapted from Woolard, 2000). Participants were provided with a set of sentences and divided into groups of three. Each group had a missing node word, which participants were asked to guess based on contextual clues (Boers et al., 2014) and collocates. In order to do so, they had to first infer the meaning of the whole collocation from the context; second, identify the collocate and the part of speech it belongs to; and finally, decide what part of speech its node might be. This exercise served as both an enhancer of participants' awareness of morphosyntactic links between words and a bridge to the theme of subsequent lessons, strength of association.

Lesson 3. Lesson 3 explored the frequency of occurrence and co-occurrence. The vast majority of researchers (Boers et al., 2014; Gitsaki, 1996; Siyanova & Schmitt, 2008) have identified frequency of occurrence as a significant factor in collocation acquisition. Additionally, binary oppositions of low – high frequency collocations emerged as statistically significant score predictors at Stage II of this research. The frequency of co-occurrence –that is, the strength of association– according to Edmonds and Gudmestad (2014), Ellis et al. (2008), and Nguyen and Webb (2016), is particularly difficult to grasp for L2 learners. This was also confirmed at Stage II of this research, where collocations with higher MI scores proved to be more problematic for participants. The following exercises were included in this lesson:

218

Finding collocates for the node (adapted from Wolter and Gyllstad, 2013). In this exercise, participants were asked to review the collocates they selected from the list of words from Lesson 2 and think about which of them were more and less frequent. Next, participants referred to the online collocation databases (for example, Just-The-Word) to check their initial guess and determine the frequency score for each collocation. While discussing the mismatch between their original judgement and the frequency score in the database, participants were introduced to the concept of corpora and explored sample sentences in the databases.

Inferring from context (adapted from Schmitt, 2000). To consolidate the concept of low and high frequency collocations and raise their awareness of the semantic and morphosyntactic links between the words, participants were given a list of collocates and asked to provide an appropriate node for each collocate. Unlike in previous exercises, this time they had to refrain from using dictionaries or online collocation databases to trigger their intuitive memory. After they had come up with correct nodes, participants were asked to determine which of these collocations were more or less frequent. This discussion helped them understand the principle of weaker and stronger collocations that allow for more and less variability, respectively.

Corpus analysis and matching (partially adapted from Jiang, 2009). This exercise offered a set of collocations and a corresponding set of MI scores in two columns and required participants to decide which collocations were weaker or stronger by matching collocations and their MI scores without referring to any additional resources.

Analyzing associative relations (adapted from McCarthy and O'Dell, 2005). This exercise focused on both the frequency of occurrence and co-occurrence since it asked

participants to think of as many collocates as they could for a range of node words (using Pro-Writing Aid: Free Online Collocation Dictionary), compare their frequencies, and then determine their strength of association based on the number of collocates for each node. This exercise led to a discussion of stronger collocations, in which, as a rule, nodes collocate with fewer dependent words.

Lesson 4. Lesson 4 focused on review and consolidation of the material. The lesson's objectives targeted recognition and controlled production of the 34 whole collocations in the context of their structure, semantics, and functions. Consequently, learning tasks transitioned from rigid sentence-level analysis-based exercises to more flexible tasks emulating real-life academic writing, revising, and speaking. The following activities were included in this lesson:

Correcting synonyms (modelled after Hill et al., 2000; Jafarpour et al., 2013; and Webb and Kagimoto, 2011). This activity replicated the acceptability judgement task used at Stage I of this research. Participants were given a set of sentences retrieved from the BNC and COCA, which had a set of deliberately inserted incorrect collocations, and were asked to find and correct the mistakes. The subsequent discussion of collocation components and their synonyms helped identify the frequent problem of paraphrasing/blending collocations.

Reconstructing the text (modelled after Lackman, 2011). For homework, participants were asked to write a short story using 10 collocations that they had learned, and then retype the story leaving blanks in place of the collocations. In class, they exchanged their stories and were asked to fill in the blanks based on the story context.

This activity helped to develop participants' ability to notice, recognize, and produce collocations based on contextual clues.

Speaking (modelled after Hill et al., 2000). Participants were given a list of topics related to everyday student life and a list of collocations. In small groups, they were asked to prepare and deliver a four-minute talk using 5-7 collocations. Then, they switched groups and repeated their talk in three minutes. Finally, after one more switch, they had only two minutes for the talk. This activity increased participants' awareness of the importance of collocations in speech fluency.

Summary and Future Recommendations

Chapter 7 described the collocation-targeted teaching module that was based on the principles of a blended teaching approach (Lackman, 2011; Sonbul & Schmitt, 2013; Szudarski & Carter, 2016; Szudarski & Conklin, 2014) and focused on three collocation-specific factors, namely, the frequency of occurrence, the frequency of co-occurrence (the strength of association), and five morphosyntactic types, which emerged as more or less significant score predictors at Stage II of this study. Due to the limited amount of time and participants' general unfamiliarity with the concept of collocation and its characteristics, only a limited number of exercises could be implemented. Although these exercises used strategies from both implicit and explicit approaches to collocation teaching and attempted a step away from the traditional format of fill-in-the-blank and multiple-choice exercises, they still relied on some familiar models, such as matching, finding nodes for collocates and vice versa, and correcting errors. Consequently, most of the tasks focused on recognizing and analyzing collocations rather than actively producing them in the context of free speaking/writing. Additionally, since the focus of

the instruction was on the specific number of academic collocations and their characteristics, the context of learning tasks relied on sentences rather than paragraphs. Therefore, this teaching module might be considered introductory/exploratory rather than comprehensive; it would be useful to expand it in future research by using longer texts and more production-focused tasks to increase the quantity of the input and output.

The purpose of this exploratory module was to introduce beginner EAP learner-participants to the concept of collocation and its basic characteristics; to reduce the number of "accuracy of use" lexical collocational errors, such as paraphrasing/blending collocations (Schmitt, 2013); and to develop their skills of noticing, analyzing, recognizing, and producing 34 academic English collocations retrieved from the Academic Collocation List (ACL) by Ackermann and Chen (2013). The novelty of this teaching module is in the principle of selecting collocational items on the basis of a set of collocation-specific factors and developing learners' awareness of these factors and understanding of how words connect and function together.

Chapter 8: Stage III. Pre- and Post-instruction Test

Rationale and Design

Stage II of the research identified certain relations between the collocation recognition and controlled-production test scores and the factors of the frequency of occurrence, frequency of co-occurrence, and three out of five MSTs (Verb + Noun, Adjective + Noun, and to some extent, Verb + Adjective). However, it remains to be seen whether teaching collocations based on these factors is feasible and will lead to a significant difference in collocation acquisition. Stage III seeks to confirm or refute the hypothesis of effective teaching based on blended (explicit and implicit) approach and targeted at specific characteristics of collocations discussed by Ellis et al. (2008), Handl (2008), Lewis (2000), McCarthy & O'Dell (2005), Nguyen and Webb (2016), and Peters (2016), rather than based on general principles of vocabulary teaching (Brandl, 2008; Graves, 2006; Schmitt, 2000). For this purpose, a large-scale classroom experimental study was conducted. The study included a pre- and post-instruction collocation recognition and controlled production-based test and a two-weeks-long (30-40 minutes twice a week) collocation-targeted teaching module described in Chapter 7 and Appendix N. This chapter 8 discusses the results of the analysis of the pre- and post-instruction-based test and its implications for teaching and learning collocations.

Instruments

Tests. The collocation-recognition and controlled production-based pre- and post-instruction test replicated the test model and items used at Stage II (see Appendix K). The rationale behind the decision to keep the same test format for Stage III was the necessity to focus on collocation-specific factors only and eliminate any external factors, such as

223

different testing tasks. Similarly, it was decided to keep the same 34 collocations (see Appendices H and M) to eliminate a potentially influential factor of different test items with different normed frequency and MI scores. Consequently, the test format, modelled after Revier (2009), presented 34 sentences retrieved from the BNC and COCA. The overall context of the sentences was discipline-non-specific and focused on general aspects of student academic life and activities. Each sentence had a blank that had to be filled in with an appropriate word combination (collocation) from the three-words-choice/two (or three)-columns-long matrix. Collocations varied on the scale of low – high F and MI scores and belonged to the five MSTs. The test distractors were selected from a range of synonyms following Jafarpour et al. (2013), Koya (2005), Nizonkiza (2015), and Webb and Kagimoto (2011), which fitted the overall context of a sentence. The definitions for the test items were provided below each collocational matrix in order to avoid the factor of unknown vocabulary.

The pre-instruction test (see Appendix K) was identical to the test used at Stage II; however, the post-instruction test (see Appendix O) had a different order of test items and their distractors as well as different sentences from the BNC and COCA. These changes were introduced to avoid the factor of prior memorization of the items.

Post-test survey. The survey replicated questions used in the survey at Stage II (Appendix L). The main purpose of these questions was to collect information on participants' language learning background, age, general vocabulary learning strategies, and their assessment of the test task. It contained multiple-choice and Likert scale questions and, based on Stage II, would require approximately 15 minutes for completion.

Participants

44 participants were beginner level EAP learners, predominantly speakers of Chinese as L1 (n = 41 [93.1%]), in the age group of 18-20 (n = 38 [88.3%]). The age of onset of learning English varied from 2 to 19 (see Figure 21). Most participants started learning English at 6 years old (n = 9 [20.9%]) or 10 years old (n = 8 [18.6%]). The number of years spent learning English ranged from 1 to 16, and the average time was 8.8 years (See Figure 22).
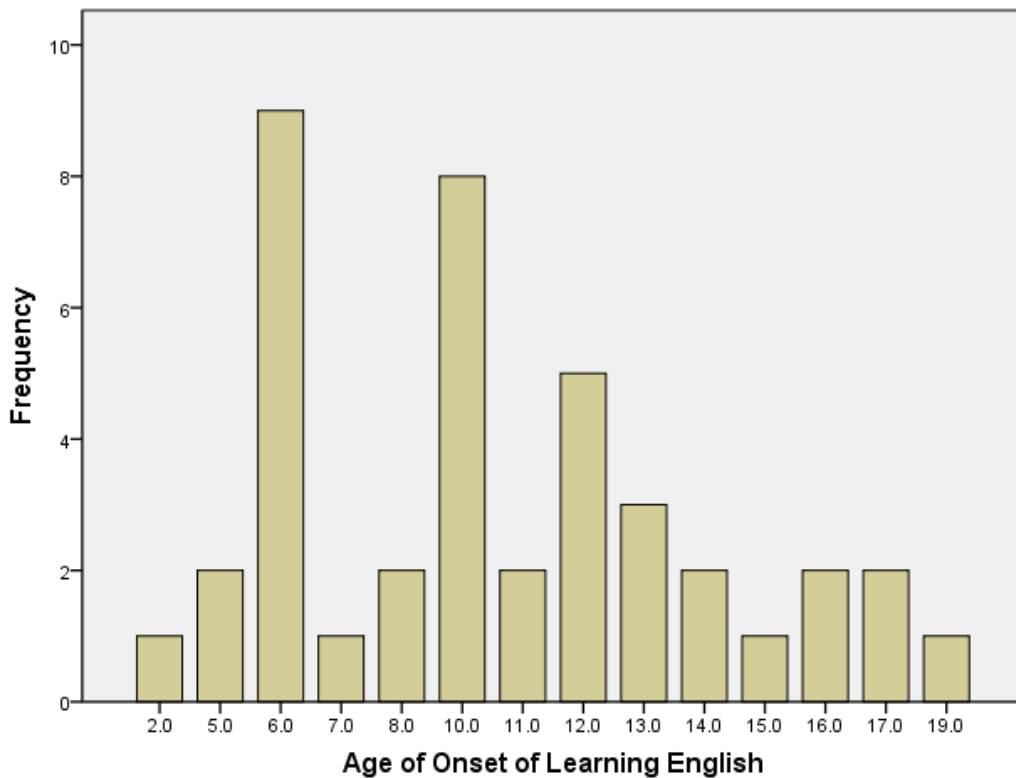
Figure 21. Age of onset of learning English.
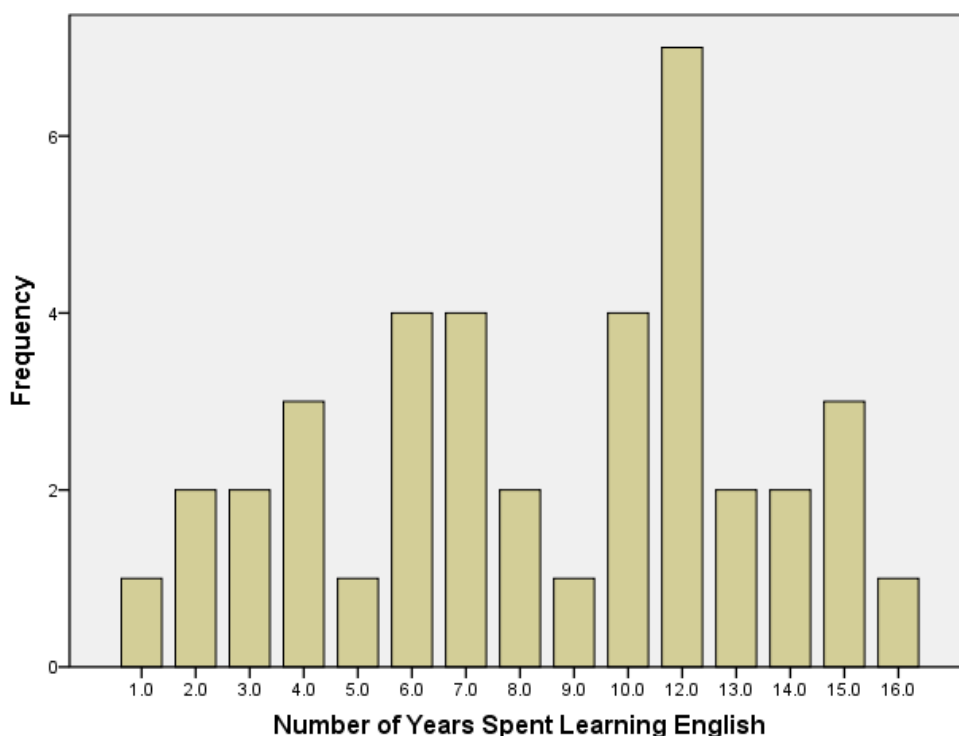
Figure 22. Number of years spent learning English.

Data Collection Procedures

The pre-instruction paper-based test was administered to participants one week

prior to instruction. Based on Stage II, the time limit of 45 minutes (approximately 1.5

minute per one test item) was considered sufficient. Replicating Szudarski and Carter

(2016) who delayed their post-test in order to measure long-term learning outcomes, the

45-minutes-long delayed post-instruction test was not administered immediately post-

instruction, but rather two weeks after the instruction ended to ensure a sufficient time

span. In order to eliminate the factor of prior memorization and elicit spontaneous

recognition and controlled production of collocations, participants were not given any

prior notice of the post-instruction test date or asked to practice and prepare for the test.

During the test, participants could work on the sentences in any order and were free to revise and correct already completed items. They were specifically asked not to dwell too long upon each sentence and to complete all questions to the best of their knowledge without leaving any missing answers. Participants were not allowed to use any external resources, such as dictionaries. After the post-instruction test, participants were asked to complete the same 15-minutes-long survey used at Stage II that collected information on participants' language background and their assessment of the test's difficulty.

Data Analysis and Results

The data was coded using SPSS Version 22.0. Both the factors that emerged as strong and weak score predictors at Stage II of the research were considered for the analysis at this stage in order to determine which types of collocations were the most and least problematic for teaching and in which areas participants demonstrated improvement. The results of the pre- and post-instruction collocation-targeted tests were analyzed in three consecutive steps.

First, descriptive statistics for categorical variables (age, L1, age of onset of learning English, and number of years spent learning English) and continuous variables (combinations of F, MI, and five MSTs) were run to ensure that there were no outliers and to compare the mean score on the two tests, and a Cronbach's alpha test was run to prove the internal consistency and validity of the test.

Second, the pre-instruction test results were compared to the results of the test administered to advanced EAP learners at Stage II of the research. An independent-samples t-test was conducted in order to identify any statistically significant differences in the overall collocation test scores between these two groups of participants (more and

less advanced EAP learners). This information was intended to shed more light on the findings of Stage I of the research, namely whether collocation knowledge depends on language proficiency level. When interpreting this data, the equal variances for the two groups were ensured with a Sig. value of greater than the cutoff of .05, the mean values of the test scores between the two groups were compared, and their statistical significance was discussed in terms of the Sig. (2-tailed) value and the cutoff of .05, t-value, and the magnitude of the differences between groups measured by the eta-squared statistics (Cohen, 1988, 1992).

Finally, the pre- and post-instruction test results in the same group of 44 beginner EAP participants were compared. The series of paired-samples t-tests were conducted to evaluate the statistically significant impact of collocation characteristics-targeted instruction on participants' collocation recognition and controlled production measured as by test scores: specifically, the overall score and respective scores in each of the categories of F, MI, five MSTs, and their combinations. While interpreting the data, the mean values of the pre-instruction test (Test 1) and post-instruction test (Test 2) were compared, and their statistical significance was proven by the Sig. (2-tailed) value less or equal to .05, the t-value of more than 0, which indicates a difference in Test 1 and Test 2 scores, and the magnitude of the teaching intervention effect as measured by the eta-squared statistics (Cohen, 1988, 1992).

Step 1. Descriptive statistics. The mean scores in each of the categories of the pre- and post-instruction test were summarized in Table 21. Since analysis indicated that 5% trimmed mean and mean values were approximately the same for each of the variables, it was concluded there were no outliers. There was only little improvement in

the pre- and post-instruction test scores (from 13.86 to 15.18, or a 6.47% increase). While most of the categories increased their average scores, there was a small decrease in the average scores for:

1. Verb + Noun collocations (2%), mostly due to the decrease in the scores for Verb + Noun collocations with low F – low MI (10%) and high F – high MI (6.67%);

2. Low F – low MI Adverb + Adjective collocations (3%);

3. High F – high MI collocations (1%), most likely because of the decrease in the scores for high F – high MI Verb + Noun (6.67%) and Verb + Adjective collocations (6%).

While the finding on the decrease in the pre- and post-instruction test scores for high F – high MI Verb + Adjective collocations is somewhat difficult to explain, the results for other two categories agree with existing research by Ellis et al. (2008) and Nguyen and Webb (2016) and suggest that ESL learners are sensitive to the frequency of occurrence, where collocations with low frequency of occurrence are the most difficult for acquisition. It also appears that certain MSTs might be more difficult for acquisition and, in contrast to Stage II of the research as described in Chapter 6, it is Verb + Noun collocations that might be least susceptible to instructional treatment. However, further analysis is needed to confirm these findings.

Table 21

Mean Scores on Pre- and Post-instruction Tests

Category	Mean scores						Percentage of increase / decrease
	Pre-instruction (Test 1)			Post-instruction (Test 2)			
	Score	Percentage in each category	Percentage out of 34	Score	Percentage in each category	Percentage out of 34	
Total Score (n = 34)	13.86	40.70	40.70	15.18	46.47	46.47	6.47
Low F – Different MI (n = 22)	8.32	37.8	24.47	9.73	44.22	28.6	6.42
Different F – High MI (n = 24)	9.5	39.58	27.9	10.45	43.54	30.7	3.96
Adjective + Noun – Different F and MI (n = 9)	3.59	39.88	10.55	3.93	43.66	11.55	3.78
Verb + Noun – Different F and MI (n = 9)	3.98	44	11.7	3.84	42	11.2	-2
Verb + Adjective –	2.39	39.83	7	2.98	49.66	8.7	9.83

| Category | Mean scores | | | | | | |
| | Pre-instruction (Test 1) | | | Post-instruction (Test 2) | | | Percentage of |
	Score	Percentage in each category	Percentage out of 34	Score	Percentage in each category	Percentage out of 34	increase / decrease
Different F and MI (n = 6)							
Verb + Adverb – Different F and MI (n = 4)	1.93	48.25	5.6	1.98	49.5	5.82	1.25
Adverb + Adjective – Different F and MI (n = 6)	1.98	33	5.8	2.45	40.8	7.2	7.8
High F – High MI (n = 12)	5.55	46	16.32	5.45	45	16.02	-1
Low F – Low MI (n = 10)	4.36	43.6	12.8	4.75	47.5	13.97	3.9
Low F – High MI (n = 12)	3.95	32.9	11.61	4.98	41.5	14.64	8.6
Adjective + Noun - High F – High MI (n = 3)	1.23	41	3.6	1.23	41	3.6	0
Adjective + Noun - Low F – Low MI (n = 3)	1.55	51.6	4.5	1.75	58.33	5.1	6.73

Category	Mean scores						Percentage of increase / decrease
	Pre-instruction (Test 1)			Post-instruction (Test 2)			
	Score	Percentage in each category	Percentage out of 34	Score	Percentage in each category	Percentage out of 34	
Adjective + Noun - Low F – High MI (n = 3)	.82	27.3	2.4	.95	31.66	2.79	4.36
Adjective + Noun - Low F – Different MI (n = 6)	2.36	39.33	6.9	2.7	45	7.94	5.67
Adjective + Noun - Different F – High MI (n = 6)	2.05	34.16	6	2.18	36.33	6.4	2.17
Verb + Noun - High F – High MI (n = 3)	1.86	62	5.47	1.66	55.33	4.88	-6.67
Verb + Noun - Low F – Low MI (n = 3)	1.14	38	3	.86	28	2	-10
Verb + Noun - Low F –	.98	32.66	2.88	1.32	44	3.88	11.34

Category	Mean scores						Percentage of increase / decrease
	Pre-instruction (Test 1)			Post-instruction (Test 2)			
	Score	Percentage in each category	Percentage out of 34	Score	Percentage in each category	Percentage out of 34	
High MI (n = 3)							
Verb + Noun - Low F – Different MI (n = 6)	2.11	35.16	6.2	2.25	37.5	6.6	2.34
Verb + Noun - Different F – High MI (n = 6)	2.84	47.33	8.35	2.91	48.5	8.55	1.17
Verb + Adjective – High F – High MI (n = 2)	.68	34	2	.75	37.5	2.2	3.5
Verb + Adjective – Low F – Low MI (n = 2)	.91	45.5	2.67	1.39	69.5	4	24
Verb + Adjective – Low F – High MI (n = 2)	.8	40	2.35	.84	42	2.47	2
Verb + Adjective – Low F	1.7	42.5	5	2.23	55.75	6.5	13.25

Category	Mean scores						Percentage of
	Pre-instruction (Test 1)			Post-instruction (Test 2)			increase /
	Score	Percentage in each category	Percentage out of 34	Score	Percentage in each category	Percentage out of 34	decrease
– Different MI (n = 4)							
Verb + Adjective – Different F – High MI (n = 2)	1.48	74	4.35	1.59	79.5	4.67	5.5
Verb + Adverb – High F – High MI (n = 2)	1.16	58	3.4	1.05	52	3	-6
Verb + Adverb – Low F – High MI (n = 2)	.77	38.5	2.2	.93	46.5	2.73	8
Adverb + Adjective – High F – High MI (n = 2)	.61	30.5	1.79	.77	38.5	2.26	8
Adverb + Adjective – Low F – Low MI (n = 2)	.8	40	2.35	.75	37	2.2	-3

234

Category	Mean scores						Percentage of increase / decrease
	Pre-instruction (Test 1)			Post-instruction (Test 2)			
	Score	Percentage in each category	Percentage out of 34	Score	Percentage in each category	Percentage out of 34	
Adverb + Adjective – Low F – High MI (n = 2)	.57	28.5	1.67	.93	46.5	2.73	18
Adverb + Adjective – Low F – Different MI (n = 4)	1.36	34	4	1.68	42	4.9	8
Adverb + Adjective – Different F – High MI (n = 4)	1.18	29.5	3.47	1.7	42.5	5	3

Pre-instruction test (Test 1). The Cronbach's alpha coefficient ($\alpha = .92$) was above the cutoff of .70, which shows a high level of internal consistency and, consequently, reliability of the test. In the corrected item-total correlation, the values of the following items were below the minimum threshold of .30:

1. Low F – high MI Verb + Adverb collocations ($\alpha = .20$);

2. Low F – low MI Verb + Adjective collocations ($\alpha = .22$);

3. High F – high MI Adverb + Adjective collocations ($\alpha = .27$);

4. Low F – different MI Adjective+ Noun collocations ($\alpha = .29$);

5. Low F – high MI Verb + Adjective collocations ($\alpha = .30$).

However, for these items the Cronbach's alpha coefficients for when the items are deleted were the same as the original alpha coefficients, which indicates that they are significant items; as such, they were included in subsequent analysis.

The Cronbach's alpha coefficient for when the item is deleted was higher than the original coefficient in case of low F – high MI Adjective + Noun collocations ($\alpha = .97$), and in the corrected item-total correlation the value was below the minimum cutoff of .30 (.-14), which indicates that this variable might be measuring something different from the scale as a whole and thus was not included in the subsequent analysis.

After removing low F – high MI Adjective + Noun collocations and then low F – high MI Verb + Adverb collocations, which had the lowest corrected item-total correlation value, neither of the remaining variables' Cronbach's alpha coefficient values exceeded the original alpha coefficient value ($\alpha = .93$), which confirms the internal consistency and reliability of the pre-instruction test (see Table 22). As a result, the test scores are consistent across the variables.

Table 22

Pre-instruction Test. Corrected Item-total Correlation and Cronbach's Alpha Coefficient Values for When the Item is Deleted

	Corrected Item – Total Correlation	Cronbach's Alpha if Item Deleted
Number of correct collocations (n = 34)	.994	.920
Low F – Different MI - All MSTs (n = 22)	.888	.918
Different F – High MI - All MSTs (n = 24)	.888	.919
Adjective + Noun – Different F – Different MI (n = 9)	.498	.925
Verb + Noun – Different F – Different MI (n = 9)	.685	.922
Verb + Adjective – Different F – Different MI (n = 6)	.526	.924
Verb + Adverb – Different F – Different MI (n = 4)	.375	.926
Adv + Adjective – Different F – Different MI (n = 6)	.598	.923
High F – High MI - All MSTs (n = 12)	.806	.920

	Corrected Item – Total Correlation	Cronbach's Alpha if Item Deleted
Low F – Low MI - All MSTs (n = 10)	.598	.923
Low F – High MI - All MSTs (n = 12)	.646	.923
Adjective + Noun – High F – High MI (n = 3)	.518	.925
Adjective + Noun – Low F – Low MI (n = 3)	.445	.925
Adjective + Noun – Low F – Different MI (n = 6)	.284	.927
Adjective + Noun – Different F – High MI (n = 6)	.363	.926
Verb + Noun – High F – High MI (n = 3)	.526	.925
Verb + Noun – Low F – Low MI (n = 3)	.368	.926
Verb + Noun – Low F – High MI (n = 3)	.585	.924
Verb + Noun – Low F – Different MI (n = 6)	.617	.923

	Corrected Item – Total Correlation	Cronbach's Alpha if Item Deleted
Verb + Noun – Different F – High MI (n = 6)	.704	.922
Verb + Adjective – High F – High MI (n = 2)	.444	.926
Verb + Adjective – Low F – Low MI (n = 2)	.213	.927
Verb + Adjective – Low F – High MI (n = 2)	.299	.927
Verb + Adjective – Low F – Different MI (n = 4)	.376	.926
Verb + Adjective – Different F – High MI (n = 4)	.479	.925
Verb + Adverb – High F – High MI (n = 2)	.401	.926
Adverb + Adjective – High F – High MI (n = 2)	.272	.927
Adverb + Adjective – Low F – Low MI (n = 2)	.372	.926
Adverb + Adjective – Low F – High MI (n = 2)	.541	.925

	Corrected Item – Total Correlation	Cronbach's Alpha if Item Deleted
Adverb + Adjective – Low F – Different MI (n = 4)	.601	.924
Adverb + Adjective – Different F – High MI (n = 4)	.535	.925

Post-instruction test (Test 2). The Cronbach's alpha coefficient (α = .94) was

above the cutoff of .70 which indicates a slightly higher level of internal consistency and

higher reliability of the post-instruction test in comparison to the pre-instruction test.

Cronbach's alpha coefficients for when the items are deleted were the same as the

original alpha coefficients for the following variables:

1. Low F – different MI Adjective + Noun collocations (α = .94);

2. Low F – low MI Verb + Noun collocations (α = .94);

3. All three categories (high F – high MI; low F – low MI; and low F – high MI) of
   Verb + Adjective collocations (α = .94);

4. Low F – low MI Adverb + Adjective collocations (α = .94);

5. Low F – high MI Adverb + Adjective collocations (α = .94).

Their values in the corrected item-total correlation were above the minimum threshold of

.30; therefore, these items appeared significant and were kept for the subsequent analysis.

The Cronbach's alpha coefficient for when the item is deleted was again higher

than the original coefficient in case of the low F – high MI Adjective + Noun collocations

(α = .94), and in the corrected item-total correlation, the value was below the minimum

cutoff of .30 (.-03), which confirms that the variable of low F – high MI Adjective +

Noun collocations was not necessary in the subsequent analysis. Additionally,

Cronbach's alpha coefficients for when the items are deleted emerged higher than the

original coefficients for the high F – high MI Verb + Adverb collocations (α = .94) and

the high F – high MI Verb + Adjective collocations (α = .94), and their values in the

corrected item-total correlation were lower than .30 (.24 and .25 respectively). As a

result, the variables of low F – high MI Adjective + Noun, high F – high MI Verb + Adverb, and high F – high MI Verb + Adjective collocations were eliminated.

When running the Cronbach's alpha coefficient analysis again without the deleted variables, none of the Cronbach's alpha coefficient values for when the item is deleted exceeded the original coefficient value of .94, and the values in the corrected item-total correlation were above the cutoff of .3, which shows a high degree of internal consistency and reliability of the post-instruction test. This means that the test scores are consistent across variables. The results are summarized in Table 23.

Step 2. Comparing test scores across language proficiency levels. An independent samples t-test compared test scores of Group 1, the advanced EAP learners from Stage II of the research, and pre-instruction test scores of Group 2, the beginner EAP learners from Stage III of the research. The mean scores for the two groups were approximately the same. The Sig. value ($p = .230$) was above the cutoff of .05, which means that the variances for the two groups were equal; however, the Sig. (2-tailed) value was .900, which indicates that there was no statistically significant difference in the test scores between Group 1 ($M = 13.75$, $SD = 4.95$) and Group 2 ($M = 13.86$, $SD = 4.08$), t (120) = -.12, $p = .900$. The magnitude of the differences in the means between the two groups was non-existent (eta-squared statistics = .000), which means that variance in the test scores cannot be explained by the differences between the groups, and consequently, language proficiency level does not seem to play a significant role in participants' performance on this test. This finding contradicts some studies (Gitsaki, 1999; Groom, 2009; Laufer & Waldman, 2011; Nizonkiza, 2012, 2015) that claim that the development of collocational competence increases as proficiency improves.

Table 23

Post-instruction Test. Corrected Item-total Correlation and Cronbach's Alpha Coefficient Values for When the Item is Deleted

	Corrected Item – Total Correlation	Cronbach's Alpha if Item Deleted
Number of correct collocations (n = 34)	.995	.942
Low F – Different MI – All MSTs (n = 22)	.931	.937
Different F – High MI – All MSTs (n = 24)	.938	.937
Adjective + Noun – Different F – Different MI (n = 9)	.677	.940
Verb + Noun – Different F – Different MI (n = 9)	.789	.938
Verb + Adjective – Different F – Different MI (n = 6)	.664	.940
Verb + Adverb – Different F – Different MI (n = 4)	.532	.942
Adverb + Adjective – Different F – Different MI (n = 6)	.683	.940

	Corrected Item – Total Correlation	Cronbach's Alpha if Item Deleted
High F – High MI – All MSTs (n = 12)	.758	.939
Low F – Low MI – All MSTs (n = 10)	.775	.938
Low F – High MI – All MSTs (n = 12)	.816	.938
Adjective + Noun – High F – High MI (n = 3)	.632	.941
Adjective + Noun – Low F – Low MI (n = 3)	.525	.942
Adjective + Noun – Low F – different MI (n = 6)	.400	.942
Adjective + Noun – Different F – High MI (n = 6)	.532	.941
Verb + Noun – High F – High MI (n = 3)	.476	.942
Verb + Noun – Low F – Low MI (n = 3)	.472	.942
Verb + Noun – Low F – High MI (n = 3)	.716	.941

	Corrected Item – Total Correlation	Cronbach's Alpha if Item Deleted
Verb + Noun – Low F – different MI (n = 6)	.797	.939
Verb + Noun – Different F – High MI (n = 6)	.667	.940
Verb + Adjective – Low F – Low MI (n = 2)	.517	.942
Verb + Adjective – Low F – High MI (n = 2)	.330	.943
Verb + Adjective – Low F – Different MI (n = 4)	.596	.941
Verb + Adjective – Different F – High MI (n = 4)	.479	.942
Verb + Adverb – Low F – High MI (n = 2)	.549	.942
Adverb + Adjective – Low F – Low MI (n = 2)	.482	.942
Adverb + Adjective – Low F – High MI (n = 2)	.523	.942
Adverb + Adjective – Low F – Different MI (n = 4)	.715	.941

	Corrected Item – Total Correlation	Cronbach's Alpha if Item Deleted
Adverb + Adjective – Different F – High MI (n = 4)	.579	.942
Adverb + Adjective – High F – High MI (n = 2)	.388	.943

Step 3. Comparing pre- and post-instruction test scores. A paired-samples t-test was conducted to explore the significance and impact of the teaching intervention for the group of beginner EAP learners on their test scores in each of the collocation-specific factor categories. For this purpose, the pre- and post-instruction test scores across each category were compared. Although Cronbach's alpha analysis determined that low F – high MI Adjective + Noun collocations did not appear to contribute to the internal consistency of both pre- and post-instruction tests, and the same could be stated about low F – high MI Verb + Adverb collocations in the pre-instruction test only, and high F – high MI Verb + Adverb and high F – high MI Verb + Adjective collocations in the post-instruction test only, all of these variables were kept for the paired t-test analysis. The reason for this decision is that it is important to follow any statistically significant improvement across categories of collocations, considering both weak and strong test score predictors. The results of the paired-samples t-test are summarized in Table 24.

Overall test score. As shown in Table 24, there was a statistically significant increase in test scores from Test 1 ($M = 13.86$, $SD = 4.08$) to Test 2 ($M = 15.18$, $SD = 4.98$), $t(43) = - 2.51$, p <. 016 (two-tailed). The mean increase in the overall test scores was 1.31 with a 95% confidence interval ranging from -2.37 to -.26. The eta-squared statistics (.12) indicated a large effect size. These results suggest that the teaching intervention contributed to the small improvement of collocation recognition and controlled production, which corresponds to Granena and Long (2013) and Nesselhauf (2005), who argue that collocation acquisition is a life-long process and cannot be mastered in a short time frame.

Table 24

Pre- and Post-instruction Tests. Corrected Item-total Correlation and Cronbach's Alpha Coefficient Values for When the Item is Deleted

Paired Samples Test

| Test 1 – Test 2 | Paired Differences | | | | | | | Sig. (2-tailed) |
| | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | | t | df | |
				Lower	Upper			
Pair 1: Number of correct collocations	-1.3181	3.4826	.52503	-2.37701	-.25935	-2.511	43	.016
Pair 2: Low F – Different MI – All MSTs	-1.409	2.739	.413	-2.242	-.576	-3.412	43	.001
Pair 3: Different F – High MI – All MSTs	-.955	3.118	.470	-1.902	-.007	-2.031	43	.048
Pair 4: Adjective + Noun – Different F – Different MI	-.341	1.462	.220	-.785	.104	-1.547	43	.129
Pair 5: Verb + Noun – Different F – Different MI	.136	1.693	.255	-.378	.651	.534	43	.596
Pair 6: Verb + Adjective – Different F – Different MI	-.591	1.499	.226	-1.047	-.135	-2.615	43	.012

248

Paired Samples Test

| Test 1 – Test 2 | Paired Differences | | | | | | | Sig. (2-tailed) |
| | Mean | Std. Devia-tion | Std. Error Mean | 95% Confidence Interval of the Difference | | t | df | |
				Lower	Upper			
Pair 7: Verb + Adverb – Different F – Different MI	-.045	1.509	.227	-.504	.413	-.200	43	.843
Pair 8: Adverb + Adjective – Different F – Different MI	-.477	1.267	.191	-.862	-.092	-2.500	43	.016
Pair 9: High F – High MI – All MSTs	.091	1.998	.301	-.517	.698	.302	43	.764
Pair 10: Low F – Low MI – All MSTs	-.386	1.883	.284	-.959	.186	-1.361	43	.181
Pair 11: Low F – High MI – All MSTs	-1.023	2.118	.319	-1.667	-.379	-3.202	43	.003
Pair 12: Adjective + Noun – High F – High MI	.000	1.034	.156	-.314	.314	.000	43	1.000
Pair 13: Adjective + Noun – Low F – Low MI	-.205	.930	.140	-.487	.078	-1.460	43	.152
Pair 14: Adjective + Noun – Low F – High MI	-.136	.878	.132	-.403	.131	-1.030	43	.309

Paired Samples Test

Test 1 – Test 2	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 15: Adjective + Noun – Low F – Different MI	-.341	1.140	.172	-.688	.006	-1.984	43	.054
Pair 16: Adjective + Noun – Different F – High MI	-.136	1.268	.191	-.522	.249	-.713	43	.480
Pair 17: Verb + Noun – High F – High MI	.205	.904	.136	-.070	.479	1.500	43	.141
Pair 18: Verb + Noun – Low F – Low MI	.273	1.020	.154	-.037	.583	1.774	43	.083
Pair 19: Verb + Noun – Low F – High MI	-.341	.987	.149	-.641	-.041	-2.291	43	.027
Pair 20: Verb + Noun – Low F – Different MI	-.136	1.472	.222	-.584	.311	-.614	43	.542
Pair 21: Verb + Noun – Different F – High MI	-.068	1.336	.201	-.474	.338	-.338	43	.737
Pair 22: Verb + Adjective – High F – High MI	-.068	.950	.143	-.357	.221	-.476	43	.636

Paired Samples Test

Test 1 – Test 2	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 23: Verb + Adjective – Low F – Low MI	-.477	.952	.144	-.767	-.188	-3.325	43	.002
Pair 24: Verb + Adjective – Low F – High MI	-.045	.914	.138	-.323	.232	-.330	43	.743
Pair 25: Verb + Adjective – Low F – Different MI	-.523	1.131	.170	-.866	-.179	-3.067	43	.004
Pair 26: Verb + Adjective – Different F – High MI	-.114	1.385	.209	-.535	.307	-.544	43	.589
Pair 27: Verb + Adverb – High F – High MI	.114	.945	.143	-.174	.401	.797	43	.430
Pair 28: Verb + Adverb – Low F – High MI	-.159	1.033	.156	-.473	.155	-1.022	43	.313
Pair 29: Adverb + Adjective – High F – High MI	-.159	.776	.117	-.395	.077	-1.360	43	.181
Pair 30: Adverb + Adjective – Low F – Low MI	.045	.714	.108	-.172	.262	.422	43	.675

Paired Samples Test

| Test 1 – Test 2 | Paired Differences | | | | | | | Sig. (2-tailed) |
| | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | | t | df | |
				Lower	Upper			
Pair 31: Adverb + Adjective – Low F – High MI	-.364	.838	.126	-.618	-.109	-2.879	43	.006
Pair 32: Adverb + Adjective – Low F – Different MI	-.318	1.073	.162	-.645	.008	-1.966	43	.056
Pair 33: Adverb + Adjective – Different F – High MI	-.523	1.045	.158	-.841	-.205	-3.317	43	.002

Test scores on five MSTs. Based on the results shown in Table 24, only two of the five MSTs had a statistically significant increase in test scores: for Verb + Adjective collocations with different F and MI scores from Test 1 ($M = 2.39$, $SD = 1.29$) to Test 2 ($M = 2.98$, $SD = 1.24$), $t(43) = -2.61$, $p = .012$; and for Adverb + Adjective collocations with different F and MI scores from Test 1 ($M = 1.98$, $SD = 1.28$) to Test 2 ($M = 2.45$, $SD = 1.21$), $t(43) = -2.50$, $p = .016$. The mean increase in these two respective scores for two MSTs was .59 for Verb + Adjective collocations with a 95% confidence interval ranging from 1.04 to 1.35, and .47 for Adverb + Adjective collocations with a 95% confidence interval ranging from .86 to .09. The eta-squared statistics for Verb + Adjective (.13) and Adverb + Adjective (.12) indicated a large effect size. This finding suggests that Verb + Adjective and Adverb + Adjective collocations were most susceptible to teaching intervention. This result is not reflected in the existing literature, possibly because these MSTs are generally overlooked in studies on collocation acquisition that compare different MSTs. However, as per Almela (2011), Erman (2014), and Granger (1998), Verb + Adjective and Adverb + Adjective collocations are usually moderately difficult for acquisition due to a certain degree of variability.

It is worth mentioning that at Stage II of the research, the Verb + Adjective type emerged as a moderately strong score predictor of test scores: the beta value of Verb + Adjective collocations was .26, and the PPMCC was .62. On the contrary, the Adverb + Adjective type did not emerge as significant at Stage II. Although the beta value of Adverb + Adjective collocations was .29 and approached the cutoff of .3, the PPMCC was even lower ($r = .60$). These results demonstrate that the collocation-targeted teaching

intervention was effective for one moderately strong score predictor (Verb + Adjective collocations) and one weak score predictor (Adverb + Adjective collocations).

The differences between Verb + Noun collocation scores with different F and MI scores did not emerge as statistically significant from Test 1 ($M = 3.98$, $SD = 1.79$) to Test 2 ($M = 3.84$, $SD = 1.95$), $t(43) = .53$, $p = .596$. This means that the decrease in pre- and post-instruction test scores identified in the descriptive statistics cannot be explained by the teaching intervention, and that other factors that are not accounted for in this study, such as semantics of the core verb (Nesselhauf, 2005), might be at play.

The differences between Adjective + Noun collocation scores with different F and MI scores also did not appear statistically significant from Test 1 ($M = 3.59$, $SD = 1.43$) to Test 2 ($M = 3.93$, $SD = 1.54$), $t(43) = -1.54$, $p = .129$, which means that the teaching intervention did not influence the increase in the test scores either. It is also important to remember that at the stage of the Cronbach's alpha coefficient analysis, low F – High MI Adjective + Noun collocations appeared as not measuring the impact of teaching due to their high alpha coefficients for when the item is deleted values. This might be the cause for the overall insignificance of this MST in spite of the fact that it emerged as a strong score predictor at Stage II.

The test scores for Verb + Adverb collocations with different F and MI scores, which emerged as weak score predictors at Stage II of the research, were also not statistically significant from Test 1 ($M = 1.93$, $SD = 1.12$) to Test 2 ($M = 1.98$, $SD = 1.06$), $t(43) = -.2$, $p = .843$. This might be due to the fact that their Cronbach's alpha coefficients for when the items are deleted values were too high (in the case of low F – high MI Verb + Adverb collocations in the pre-instruction test and high F – high MI Verb

+ Adverb collocations in the post-instruction test), which means that the scores were not consistent and these variables could have been eliminated from the analysis. Therefore, we cannot predict the impact of teaching intervention upon this MST.

On the whole, it appears that collocation-targeted instruction did not make a statistically significant impact on the strong score predictors, that is, Verb + Noun and Adjective + Noun collocations. However, it was effective in the case of two other MSTs, that is, Verb +Adjective and Adverb + Adjective collocations. It is worth pointing out that Verb + Adjective collocations also emerged as moderately strong score predictors at Stage II of the research. Therefore, it appears that more teaching focus on this MST, which is frequently not addressed in the research, might contribute to collocation acquisition.

Test scores on the categories of F and MI. The three F-MI scale categories had a statistically significant increase in test scores:

1. Low F – different MI collocations, from Test 1 ($M = 8.32$, $SD = 2.54$) to Test 2 ($M = 9.73$, $SD = 3.52$), $t(43) = - 3.41$, $p = .001$;

2. High MI – different F collocations, from Test 1 ($M = 9.5$, $SD = 3.17$) to Test 2 ($M = 10.45$, $SD = 3.53$), $t(43) = -2.03$, $p = .048$;

3. Low F – high MI collocations, from Test 1 ($M = 3.95$, $SD = 1.65$) to Test 2 ($M = 4.98$, $SD = 2$), $t(43) = -3.20$, $p = .003$.

The mean increase in these three scores was 1.4 for low F – different MI collocations with a 95% confidence interval ranging from -2.24 to -.57; .95 for high MI – different F collocations with a 95% confidence interval ranging from -1.90 to -.007; and 1.02 for low F – high MI collocations with a 95% confidence interval ranging from -1.66

to -.37. The eta- squared statistics for low F – different MI collocations (.21) and low F – high MI collocations (.19) indicated a large effect size, and the eta-squared statistics for high MI – different F collocations (.08) indicated a moderate effect size. This result signifies the statistical significance of the increase in the test scores for collocations with low frequency of occurrence, which, according to a number of studies (Granger & Bestgen, 2014; Siyanova-Chanturia & Spina, 2015; Szudarski & Conklin, 2014; Wolter & Gyllstad, 2013), were found to be particularly problematic for ESL learners. Therefore, it appears that collocation-targeted instruction was effective while increasing learners' awareness of less frequent collocations. However, the factor of the frequency of co-occurrence might interact with this result, since the largest statistically significant improvement was noticed for collocations with high MI scores.

For the category of low F – low MI collocations, the difference in the test scores was not statistically significant from Test 1 ($M = 4.36$, $SD = 1.83$) to Test 2 ($M = 4.75$, $SD = 2.03$), $t(43) = -1.36$, $p = .181$. The same can be stated about high F – high MI collocations, from Test 1 ($M = 5.55$, $SD = 2.11$) to Test 2 ($M = 5.45$, $SD = 2.09$), $t(43) = .302$, $p = .764$. This corresponds to the results of descriptive statistics and the Cronbach's alpha coefficient analysis that determined certain high F – high MI collocation types as not impacted by teaching intervention. This new finding confirms the hypothesis of the interrelatedness between low frequency of occurrence and high frequency of co-occurrence as summarized in the previous paragraph, and consequently, the variables of frequency of occurrence and co-occurrence are interconnected as per Stage II of the research and other research studies by Barnbrook et al. (2013), Durrant and Schmitt (2009), and Siyanova and Schmitt (2008).

Overall, these findings suggest that there might be a positive impact of teaching intervention on low F - high MI collocations, but not on high F – high MI and low F – low MI collocations. Since low F – high MI collocations represent the strongest and less variable collocational types, identified as particularly difficult for the ESL learners in studies by Siyanova and Schmitt (2008) and Nguyen and Webb (2016), it appears that the instruction targeted at collocational characteristics might be effective indeed.

Test scores on the categories of low F – high MI for five MSTs. In order to examine how the binary opposition of low F – high MI that emerged as statistically significant might have interacted with MSTs, the results of the paired-samples t-test (see Table 24) were analyzed for the binary opposition of low F – high MI across five MSTs.

The results for low F – high MI Verb + Adjective collocations did not demonstrate a statistically significant increase from Test 1 ($M = .80$, $SD = .73$) to Test 2 ($M = .84$, $SD = .64$), $t(43) = -.33$, $p = .743$. This means that although the MST of Verb + Adjective collocations demonstrated a significant improvement in post-instruction test scores, the particular category of low F – high MI Verb + Adjective collocations was not impacted by the teaching intervention. However, when examining this MST across the binary oppositions of F and MI, we can see that low F – low MI Verb + Adjective collocations had the largest statistically significant increase from Test 1 ($M = .91$, $SD = .67$) to Test 2 ($M = 1.39$, $SD = .68$), $t(43) = -3.32$, $p = .002$. The mean increase was -.47 with a 95% confidence interval ranging from -.76 to .18. The eta squared statistic (.20) indicated a large effect size. This indicates that this variable could have been impacted by the teaching intervention, but because low F – low MI collocations –regardless of their

MST– were found as statistically insignificant ($p = .181$), this result could be dismissed as an anomaly.

On the contrary, low F – high MI Adverb + Adjective collocations demonstrated a statistically significant score increase from Test 1 ($M = .57$, $SD = .58$) to Test 2 ($M = .93$, $SD = .69$), $t(43) = -2.87$, $p = .006$. The mean increase was -.36 with a 95% confidence interval ranging from -.61 to -.10. The eta-squared statistics (.16) indicated a large effect size. This result indicates that Adverb + Adjective collocations with low F and high MI scores were susceptible to the teaching treatment.

The other two MSTs – Adjective + Noun collocations with low F and high MI scores ($p = .309$) and Verb + Adverb collocations with low F and high MI scores ($p = .313$) predictably did not show any statistically significant improvement because their Sig. 2-tailed values were more than the maximum threshold of .05. They also did not show any statistically significant improvement in overall post-instruction test scores, as discussed at previous stages of analysis.

However, even though the MST of Verb + Noun collocations did not show any improvement in post-instruction test scores, the category of low F – high MI Verb + Noun collocations demonstrated a statistically significant score increase from Test 1 ($M = .98$, $SD = .85$) to Test 2 ($M = 1.32$, $SD = .88$), $t(43) = -2.29$, $p = .027$. The mean increase was -.34 with a 95% confidence interval ranging from -.64 to -.04. The eta-squared statistics (.10) indicated a moderate effect size. This result signifies that although there was a statistically non-significant decrease in post-instruction test scores for Verb + Noun collocations, one particular category of Verb + Noun collocations with low F and high MI scores was susceptible to teaching.

The overall results of this phase of analysis confirmed that low F – high MI collocations were generally more susceptible to treatment in comparison to other binary oppositions of F and MI for two out of five MSTs.

Post-test survey. The first part of the post-test survey focused on participants' first and second languages, as well as age and the amount of time they spent learning English in various academic and non-academic settings. Some of these variables are considered valid external factors influencing acquisition of formulaic sequences –for example, the age of onset of learning English (Granena & Long, 2013; Wray, 2002), the amount of time spent learning English (Wolter & Gyllstad, 2013), and L1/L2 interference (Carol et al., 2016; Millar, 2011; Yamashita & Jiang, 2010). However, the data collected at Stage III did not provide enough variability, that is, most participants shared the same L1 (Chinese), were in the same age group (18-20 years old), and had spent approximately the same number of years of learning English. Therefore, these results were not statistically significant in terms of their influence on the test scores. Consequently, they were not used for data analysis and mostly served to provide background information on participants as described in the above section "Participants".

The second part of the survey collected information on participants' preferred word-learning strategies (one-by-one, in combinations/patterns, or "other"). Over 65% of participants indicated that their preferred strategy was learning words in combination; however, the results of a one-way ANOVA confirmed that this did not seem to impact their pre- or post-instruction test scores in terms of statistical significance. Further, those participants who selected the "learning words one-by-one" option did not necessarily have lower test scores. This finding contradicts the findings of Stage I of the research,

which identified vocabulary learning strategies as a potentially influential factor, and of previous studies by Lewis (2000), Mian (1988), and Webb and Kagimoto (2009). The absence of a statistically significant impact can be explained by the interaction of this factor with other external factors, such as L1, age, and the number of years spent learning English, for which the data did not provide enough variability. Another explanation might be that participants were used to learning words/word combinations by heart, that is, memorizing/learning them intrinsically, which according to Lackman (2011), McCarthy and O'Dell (2005), Szudarski and Carter (2016), Szudarski and Conklin (2014), and Woolard (2005) might not be the most effective approach and needs to be combined with other explicit vocabulary learning strategies.

The last question of the survey asked participants to rate the post-instruction test in terms of its difficulty on the scale from 1 ("very easy") to 5 ("very difficult"). Most participants rated the test at 3 ("somewhat difficult") or 4 ("difficult"). There was no significant difference in the test assessment between Stage II participants for whom this task was completely new and Stage III participants for whom the task was not new because they had already completed the pre-instruction test. It appears that the test assessment by the participants does not depend on the familiarity of the task, but rather on its difficulty, as well as the difficulty of and familiarity with the test items.

Discussion

The data from the pre- and post-instruction collocation-targeted tests, which replicated the test used for Stage II of this research and was modelled after Revier (2009), were analyzed for a statistically significant change in the recognition and controlled production of the 34 academic English collocations retrieved from the Academic

Collocation List (ACL) (Ackermann & Chen, 2013) as a result of the four lessons of 30-40-minutes each over two weeks, where the instruction focused on collocation characteristics such as frequency of occurrence, frequency of co-occurrence, and morphosyntactic types. A very small statistically significant improvement in test scores (6.47%) was identified. This corresponds to findings from previous studies by Bardovi-Harlig and Bastos (2011), Cieślicka (2015), Granena & Long (2013), Nesselhauf (2005), Woolard (2001), and Wray (2002), who indicated that the development of collocational competence is a long process that cannot be mastered in a short time frame. It is important to note, nevertheless, that participants had already spent an average of 10 years learning English, and, according to the results of the post-test survey, over 65% of participants were familiar with learning/memorizing words in patterns or combinations, which emerged as a significant external factor at Stage I of the research and was discussed in other studies by Lewis (2000), McCarthy & O'Dell (2005), Webb and Kagimoto (2009), Webb et al. (2012), and Wray (2002). It appears, therefore, that simply memorizing words in combinations, that is, learning them intrinsically, might not be sufficient for mastering collocations, and explicit vocabulary learning tasks targeted at noticing, active engagement, and deep processing could also be useful (Ellis, 1995; Lewis, 2000; Nesselhauf, 2005; Schmitt, 2010; Szudarski & Carter, 2016).

In order to account for the influence of another external factor, namely language proficiency, which emerged as significant at Stage I of the research, an independent-samples t-test was run for the test scores of Group 1 (more advanced EAP learners from Stage II) and Group 2 (beginner EAP learners from Stage III, where only pre-instruction test results were considered). The findings demonstrated that the difference between the

levels was small and not statistically significant, and therefore, level of language proficiency cannot adequately account for the differences in test scores, which contradicts previous research studies (Cieślicka, 2015; Edmonds & Gudmestad, 2014; Li & Schmitt, 2010). One possible explanation for this finding might be that the difference in language proficiency between the two groups in this study might simply not be large enough because their age, primary language, age of onset of learning English, and number of years spent learning English were approximately the same. Yet another factor to consider is that participants in both groups were relatively new to the concept of collocations and the test format, which could have potentially impacted their test scores.

While the differences between the overall pre- and post-instruction test scores in the beginner learners' group did not appear as particularly large, it was still worthwhile to explore each of the variables, that is, the binary oppositions of high F – high MI, low F – low MI, low F – high MI, and the five MSTs, and any statistically significant changes in their scores by means of descriptive statistics and a paired samples t-test.

The results of the t-test suggested that an increase in the test scores was mostly due to participants' ability to recognize and produce the binary opposition of low F – high MI collocations. It is important to mention that Stage II of the research identified that low F – high MI collocations were the hardest for participants to recognize and accounted for the smallest percentage of correct responses. This finding was in accordance with what researchers such as Ellis et al. (2008), Durrant & Schmitt (2009), Nguyen and Webb (2016), and Wolter and Gyllstad (2013) have suggested about collocations with low frequency of occurrence and high frequency of co-occurrence; specifically, that these are the most challenging for ESL learners because, according to Li

and Schmitt (2010), Nguyen and Webb (2016), and Revier (2009), infrequent

collocations with higher MI scores have a less transparent meaning, allow for less

variability, represent multiword units with the higher strength of association ties and

therefore, possess more collocation characteristics.

It appears that the collocation-targeted instruction influenced participants' ability

to recognize and produce more challenging collocations; however, it largely did not

influence their recognition and controlled production of low F – low MI collocations or

less challenging high F – high MI collocations, which did not emerge as statistically

significant in test scores or as influential on the test's internal consistency. A possible

explanation might be that possessing a less semantically transparent meaning and being

less flexible, less variable, and more strongly associated, low F – high MI collocations

are more likely to be memorized as whole units as a result of repetition-based input (each

collocation was repeated a minimum of four times across the four lessons) than their

more variable counterparts. For example, the collocation *cast doubt* (F = 1.75, MI =

10.07) is less flexible and semantically transparent and thus, upon repetition, is more

easily recognized than the collocation *play a role* (F = 14.23, MI = 9.04), which also has

a high MI score but due to its high frequency of occurrence, allows for more flexibility

(e.g., *play a part*); or the collocation *give access* (F = 1.03, MI = 3.76), which is

infrequent, weakly associated, possesses a more direct meaning, and consequently, is

closer to free (non-formulaic) word combinations, such as *give love/comfort/hope*.

Another explanation is connected to other potentially influential collocation-specific and

frequency-related factors, such as frequency of individual components (nodes or

collocates), which was identified as significant in a range of studies by Hoffmann and

Lehmann (2000), Macis and Schmitt (2017), Nguyen and Webb (2016), and Nizonkiza (2015). These factors were beyond the scope of this study, yet it might be helpful to address them in future research.

Another important finding was that participants demonstrated a statistically significant improvement for two out of five MSTs, namely Verb + Adjective and Adverb + Adjective collocations. However, according to Stage II of the research, unlike Verb + Noun collocations that accounted for 44% of correct responses, these two MSTs had a lesser percentage of correct responses: Adverb + Adjective collocations (34.5%) and Verb + Adjective collocations (38.7%). Moreover, these two MSTs had only a moderate statistical significance when correlated with other MSTs: Verb + Adjective collocations ($r = .65$) and Adverb + Adjective collocations ($r = .61$) as compared to Adjective + Noun collocations ($r = .71$), and Verb + Noun collocations ($r = .81$). Additionally, while Adjective + Noun and Verb + Noun collocations were identified as the strongest score predictors, Verb + Adjective collocations were only moderately significant score predictors, while Adverb + Adjective collocations were weak score predictors at Stage II of the research.

Adverb + Adjective and Verb + Adjective collocations are less addressed in the literature as compared to more popular Verb + Noun and Adjective + Noun collocations, and are also considered as the least characteristic and most variable of all MST types (Almela, 2011; Edmonds & Gudmestad, 2014). Notwithstanding the fact that these are not typical and highly variable collocations and seemed to present more of a challenge for L2 learners (as Stage II of this research demonstrated), collocation-characteristics-targeted instruction appears to be the most influential on the participants' recognition and

264

controlled production of these two MSTs. This finding is hard to explain because first, studies that explore how teaching these two MSTs would influence their acquisition are next to none, and second, as Stage II determined, Adverb + Adjective and Verb + Adjective collocations have average F and MI scores that are relatively similar to the scores of those MSTs that emerged as not susceptible to teaching treatment (Verb + Noun and Adjective + Noun) at this stage. This finding is even more interesting because Verb + Noun and Adjective + Noun collocations are generally considered to be more typical and less variable collocations that consequently are more frequently addressed in the literature (Boers et al., 2014; Eyckmans, 2009; Gyllstad, 2009; Jaén, 2007; Koya, 2005; Pellicer-Sánchez, 2017) and were identified as stronger score predictors at Stage II of this research.

When providing a descriptive statistics overview for the increase/decrease of pre- and post-instruction test scores across the variables representing collocations of different MSTs and F and MI scores, there was an increase for most of the variables except Verb + Noun collocations (2% of decrease in test scores). This finding is in accordance with studies by Boers et al. (2014), Eyckmans (2009), and Gyllstad (2009), which argued that Verb + Noun collocations are among the most problematic for learners and therefore, are slow to be acquired. However, this finding is in disagreement with the results of Stage II of this research, which indicated that Verb + Noun collocations as the most easily recognized, have the highest percentage of correct responses, and consequently, are most likely to be successfully mastered. The finding also did not emerge as statistically significant in the subsequent paired-samples t-test analysis. This means that we cannot confirm that the teaching intervention had a negative impact or any impact at all on Verb

265

+ Noun collocations. However, even though this morphosyntactic collocational type did not emerge as susceptible to teaching, one particular category of them, low F – high MI Verb + Noun collocations, demonstrated a statistically significant improvement in post-instruction test scores. This finding serves as additional evidence of the importance of collocation-targeted instruction for the recognition and controlled production of low F – high MI collocations.

On the whole, it appears that those collocations that were confirmed as weaker score predictors at Stage II of the research and possessed fewer collocational characteristics (Almela, 2011; Edmonds & Gudmestad, 2014), namely Verb + Adjective and Adverb + Adjective collocations, were the most susceptible to teaching intervention. At the same time, collocation characteristics-targeted instruction did not appear to have much effect on those collocations that emerged as stronger score predictors at Stage II, namely Adjective + Noun and Verb + Noun collocations (with the exception of low F – high MI Verb + Noun collocations).

The research results can be explained by the interrelatedness of not only the factors of frequency of occurrence and co-occurrence as confirmed at Stage II of the research and promoted by researchers such as Barnbrook et al. (2013), Granger and Bestgen (2014), Metin and Karaoğlan (2011), and Siyanova and Schmitt (2008), but also the factors of frequency of occurrence, co-occurrence, and morphosyntactic structure. We can observe that the largest improvement in test scores was for low F – high MI Adverb + Adjective collocations and Verb + Noun collocations. Thus, it appears that as Stage II of the research suggested, the factors of frequency of occurrence and co-occurrence might impact different morphosyntactic types to various extents. As far as Chapters 3 and 4

demonstrate, this finding has not been discussed in the literature, and therefore, more research on how frequency of occurrence and co-occurrence interacts with and manifests itself in morphosyntactic structure is needed to explain this phenomenon.

Although the overall effect of the collocation-focused instruction was not large, as the relatively small increase in the pre- and post-instruction test scores demonstrates, it is nevertheless important and can serve as a solid basis for future research studies. First, it demonstrated that collocations can be taught based on their specific characteristics, and that such treatment that blends implicit and explicit teaching approaches (Mahvelati & Mukundan, 2012; Sonbul & Schmitt, 2013; Szudarski & Carter, 2016; Szudarski & Conklin, 2014) does increase the ESL learners' awareness of different types of collocations. In addition, the results of analysis showed which types of collocations appear to be most susceptible to teaching (namely, Verb + Adjective and Adverb + Adjective collocations and, to some extent, Verb + Noun collocations with low frequency of occurrence and high frequency of co-occurrence) and which collocations seem to be more problematic (Adjective + Noun, Verb + Adverb, and, to some extent, Verb + Noun collocations with high and low frequency of occurrence and co-occurrence). Prior to this research, only a few studies (Ellis et al., 2008; Handl, 2008; Nguyen & Webb, 2016; Peters, 2016; Webb & Kagimoto, 2011) engaged in a comparison of different collocational types based on their multidimensional classifications, and, to the best of my knowledge, none have attempted to teach and then examine the results of instruction targeted at the intralinguistic (collocation-specific) factors.

Summary and Future Recommendations

Chapter 8 described the pre- and post-instruction collocation recognition and controlled production-targeted tests modelled after Revier (2009) and replicated from Stage II of this research. The purpose of these tests was to determine the effectiveness of teaching collocations (the teaching module is described in Chapter 7 and Appendix N) that focused on such collocation characteristics as frequency or occurrence and co-occurrence and five morphosyntactic types (Adverb + Adjective, Adjective + Noun, Verb + Adjective, Verb + Adverb, and Verb + Noun). The results of the statistical analysis (paired samples t-test) indicated that the improvement in terms of the pre- and post-instruction test scores was small (6.47%) but statistically significant. This improvement might have been larger were it not for the limited time available for instruction. Participants who were mostly new to the concept of collocation and learning tasks might have needed more practice to acquire and consolidate the material, since collocation competence develops slowly in L2 learners, and it might take as many as 15 repetitions (Webb et al., 2013) in order for the concept to sink in. However, it was not possible for this research to spend more than two weeks on teaching collocations.

The improvement in the test scores was mainly achieved by the increase in the average scores for such morphosyntactic types as Verb + Adjective and Adverb + Adjective. Additionally, the binary opposition of low frequency of occurrence – high frequency of co-occurrence, which manifested itself in Adverb + Adjective and Verb + Noun collocations, was identified as significant.

On the one hand, these findings were somewhat contradictory to Stage II of this research which confirmed that collocations with high F scores, in particular Verb + Noun

and Adjective + Noun collocations, were the easiest to recognize and produce and therefore the most likely to be mastered. Additionally, the results did not correspond to a range of studies by Durrant and Schmitt (2009), Ellis et al. (2008), Nguyen and Webb (2016), and Wolter and Gyllstad (2013), that claimed that collocations with low F and high MI scores took more time to learn. At the same time, the results were confirmed in other studies (Boers et al., 2014; Eyckmans, 2009; Gyllstad, 2009; Jaén, 2007; Koya, 2005; Pellicer-Sánchez, 2017) that claimed that Verb + Noun and Adjective + Noun collocations were more difficult for second language learners because of their less variable forms and less transparent meaning.

On the other hand, these results indicated that first, there might be a possible connection between the factor of morphosyntactic structure and the factor of frequency of occurrence and co-occurrence, which has not been explored in the literature so far. Second, the study's findings identified collocation types that might potentially be more and less problematic for teaching, and demonstrated that it was possible to teach collocations based on their specific characteristics by means of a blended approach based on repetition, noticing, and analyzing, as promoted in studies by Mahvelati and Mukundan (2012), Sonbul and Schmitt (2013), Szudarski and Carter (2016), and Szudarski and Conklin (2014). This might be a pathway to a range of consecutive collocation studies and teaching methodologies based on collocation-relevant factors rather than general vocabulary acquisition principles.

Some additional findings deal with the external factor that emerged as significant at Stage I of this research and in many other studies (Bardovi-Harlig & Bastos, 2011; Granger & Bestgen, 2014; Schmitt, 2013; Siyanova & Schmitt, 2008; Wang & Good,

2007; Wray, 2002, 2008), namely, language proficiency level. The results of an independent samples t-test between the two groups (EAP learners at more and less advanced levels) demonstrated that overall language proficiency level did not appear to be statistically significant and the differences between their test scores were practically non-existent. This might mean that progressing to a more advanced stage of learning English does not necessarily coincide with the development of collocational knowledge because, as previous studies (Granena & Long, 2013; Groom, 2009; Nesselhauf, 2005; Wray, 2008) have mentioned, collocational competence is a time-consuming, possibly life-long process.

Chapter 9: Conclusion and Future Directions

This chapter combines findings of three stages of the mixed-methods sequential transformative design-based research and discusses them in the context of the study objectives and research questions. Then, limitations and contributions of each of the stages are summarized, and possible directions for future research on the factors influencing collocation acquisition and collocation testing and teaching based on these factors are suggested.

Summary of Main Findings

This dissertation presents a consecutive narrative that unfolds the typology of external and internal (collocation-specific) factors that impact collocation recognition and controlled production by adult speakers of ESL and discusses how these findings can be implemented in the practice of testing and teaching collocations. Chapters 2 and 3 surveyed the history of collocational studies and the insights from the research of the past decade and suggested that although multiple studies consider different factors ranging from biological (for example, age of onset of learning a language) to psychological (for example, motivation and attitudes to language learning) and language-specific (L1 - L2 congruency), currently there is no comprehensive classification that would address multiple factors across extralinguistic, interlinguistic, and intralinguistic levels in their relation to one another. In part, this is due to the ambiguity of the concept of collocation, which has traditionally been defined in relative rather than absolute terms as a semantically semi-transparent, lexically, and morphostructurally semi-restricted formulaic sequence consisting of 2-4 adjacent or non-adjacent lexical and grammatical words and having different degrees of the frequency of occurrence and co-occurrence

271

(MI $\geq$ 3). Consequently, two related problems emerge. First, in its review of research methodologies pertaining to collocation recognition, recall, and production-focused studies, Chapter 4 noted that there is no optimal testing format that measures contextual recognition and production of whole collocations. Second, as Chapter 7 indicated, current approaches to teaching collocations are modelled after general vocabulary teaching principles rather than focused on the specific difficulties that learners experience when mastering collocations or the types of collocations that are likely to cause these difficulties.

As a result, three research questions were developed. Research Question 1 sought to identify external and internal factors impacting collocation acquisition. Research Question 2 concerned collocation characteristics as the basis for selecting collocations for teaching and testing and developing collocation-targeted learning tasks. Research Question 3 attempted to explore formats for testing recognition and/or controlled production of collocations based on collocation-specific factors. In order to answer these questions, three consecutive studies were conducted.

The first study (Stage I) described in Chapter 5 consisted of two sub-stages (a pilot study and a large-scale quantitative study) and focused on the extralinguistic and interlinguistic factors that influence recognition of false (non-English-like) collocations by adult speakers of English as L1 and L2. As per Gyllstad (2009) and Siyanova and Schmitt (2008), recognition or identifying a collocation from a set of different options and making a decision about whether this collocation is common and appropriate in English is a basic part of collocational competence. In order to identify challenges that ESL speakers might experience when recognizing collocations, it is important to explore

how speakers of English as L1 and L2 compare in their receptive knowledge. Additionally, a focus on false collocations allows to correlate participants' collocational knowledge with a number of different criteria, including extralinguistic factors such as age, gender, L1/L2, and cognitive strategies of test-taking (reading fluency, the focus of attention while reading, and attention span). For speakers of English as L2, there are additional extralinguistic factors, including length of residence in an English-speaking country, age of onset of learning English, and cognitive strategies of vocabulary learning, as well as interlinguistic factors such as English language proficiency, English as a predominant language for communication, and L1-L2 interference.

The pilot study implemented standardized cloze and multiple-choice-based and accuracy judgment tasks and semi-structured interviews with four speakers of English as L2 and one speaker of English as L1. The findings suggested that those speakers of English as L2 who implemented similar cognitive and communicative strategies to those employed by the speaker of English as L1 –that is, using mostly English for communication and reading combinations of words and focusing on both word meaning and form, as well as general "structure" and "flow" – had higher test scores. Additionally, the pilot study identified strategies that participants implemented to recognize false collocations. While some of these strategies were similar for speakers of English as L1 and L2 (for example, relying on context and background knowledge), others were specific to the ESL speakers (for example, comparing languages or mentally translating and analyzing words) and to the speaker of English as L1 (relying on intuition). These findings are in line with existing research by Siyanova (2010), Siyanova-Chanturia and Martinez (2014), and Wray (2002) that confirms that ESL speakers have a less developed

intuition, which they might compensate for by analyzing word combinations. Another important finding of this pilot study was that unlike the accuracy judgment task, the standardized multiple-choice test was not an adequate measurement of collocation recognition. All participants obtained much higher scores on this task and commented on the limited options of choice.

The second study at Stage I of the research was conducted with 50 speakers of English as L1 and 43 speakers of English as L2 and implemented an accuracy judgment task only –that is, a reading-comprehension test that was partially adapted from Siyanova and Schmitt (2008), Szudarski and Conklin (2014), and Yamashita and Jiang (2010). The test included paragraphs from Canadian newspapers with 18 deliberately inserted false Verb + Noun and Adjective + Noun collocations consisting of two adjacent lexical words. The post-test survey included questions on the participants' language and learning background and the test-taking cognitive strategies they employed. Similarly to the pilot study, the results of this stage suggested some similarities between speakers of English as L1 and those speakers of English as L2 who obtained higher test scores. Those participants who focused on word form and meaning, sentence structure, and general flow were more successful in both groups. The analysis of test scores in the group of speakers of English as L2 also identified a number of statistically significant interlinguistic factors, such as L1 and English as a predominant language for communication, and extralinguistic factors, such as learning words in combinations, self-assessment of English language proficiency, and, to some extent, the length of residence in an English-speaking country. At the same time, some extralinguistic factors, such as

age, gender, the age of onset of learning English, and cognitive test-taking strategies of reading speed and attention span while reading did not prove statistically significant.

As discussed in Chapter 6, Stage II of the research targeted intralinguistic collocation-specific factors of the frequency of occurrence (measured by the normed frequency per million) and co-occurrence (measured by the MI score) and morphosyntactic structure. The test partially modelled after Revier's (2009) CONTRIX matrix combined cloze and accuracy judgment tasks and focused on measuring recognition and controlled production of 34 whole academic English collocations that consisted of two adjacent lexical words. The test was administered in several advanced level ESL classrooms with over 90 participants.

The results of the test were in agreement with existing research by Barnbrook et al. (2013), Metin and Karaoğlan (2011), and Siyanova and Schmitt (2008) and demonstrated that the factors of frequency of occurrence and co-occurrence are highly interconnected, and dichotomies of low – high frequency of occurrence and/or co-occurrence have a significant impact on collocation recognition and controlled production. Thus, the study identified that it is easier for ESL speakers to recognize and produce high frequency collocations versus items with low frequency of occurrence. However, the situation is reversed with frequency of co-occurrence, where collocations with higher MI scores present more challenges for ESL learners. This latter finding contradicts the findings of Ellis et al. (2008), one of very few studies that compared the MI score and normed frequency as two factors influencing collocation acquisition, which did not find any negative correlation between test scores and high MI score. However, the finding of the present research follows the pattern of other studies, for example, Edmonds

and Gudmestad (2014) and Nguyen and Webb (2016), who suggest that unlike speakers of English as L1 who focus on strength of association (that is, frequency of co-occurrence), ESL speakers are predominantly guided by frequency of occurrence when learning collocations.

In terms of morphosyntactic structure, the study identified that Verb + Noun collocations were the most easily recognized by participants and also the strongest score predictors, while Adjective + Noun and Verb + Adjective collocations were close runners-up. This might be due to the fact that collocations that represent combinations of verbs and/or adjectives are more frequently used and pervasive (Almela, 2011; Nesselhauf, 2005).

A final important finding of this stage was that all three factors of morphosyntactic structure and frequency of occurrence and co-occurrence are interconnected. Thus, weaker less-frequent Adjective + Noun, Adverb + Adjective, and Verb + Adjective collocations were recognized by participants more easily than weaker non-frequent Verb + Noun collocations, while stronger more frequent Noun + Verb collocations were recognized most easily of all types.

While the first two research stages answered Research Question 1 and identified external and internal factors that impact collocation recognition and controlled production, Stage III, which is discussed in Chapters 7 and 8, focused on selecting collocational items and teaching/testing collocations based on their characteristics as explored at Stage II, namely the frequency of occurrence and co-occurrence and five morphosyntactic types. Following Schmitt (2010), Szudarski and Carter (2016), and Szudarski and Conklin (2014), the collocation-targeted teaching module represented a

combination of implicit (contextualized noticing and repetition-based) and explicit (lexical and structural analysis and active engagement-focused) exercises. These exercises were administered to 44 beginner EAP learners, and the pre- and delayed post-instruction recognition and controlled production tests were implemented. The format of the two tests and 34 collocational items were the same as the test used for Stage II of the research.

The findings of Stage III showed that the blended teaching approach and learning tasks focused on collocational characteristics were effective, as demonstrated by a 6.47% increase in post-instruction test scores, which can be considered a significant result after four lessons and two weeks only. As Granena & Long (2013), Nesselhauf (2005), and Wray (2002) confirm, the development of collocational knowledge usually takes more time. This increase in test scores was mostly due to the participants' ability to recognize and produce collocations with low frequency of occurrence and high frequency of co-occurrence, while collocations with high frequency of occurrence and co-occurrence or collocations with low frequency of occurrence and co-occurrence were processed and produced less accurately. This result is important because according to the findings of Stage II of this study and existing research by Ellis et al. (2008) and Nguyen and Webb (2016), collocations with low frequency of occurrence and high frequency of co-occurrence are less variable and transparent and exhibit more pronounced characteristics of formulaicity, and are therefore considered more difficult for ESL learners.

Regarding morphosyntactic structure, participants had higher test scores for Verb + Adjective and Adverb + Adjective collocations, which are considered more variable and transparent (Almela, 2011; Edmonds & Gudmestad, 2014). At the same time, their

knowledge of Verb + Noun and Adjective + Noun collocational types, which had emerged as strong score predictors as Stage II of the research, was not significantly affected. One exception was Verb + Noun collocations with low frequency of occurrence and high frequency of co-occurrence, which showed a statistically significant increase in post-instruction test scores. This finding confirms that frequency of occurrence and co-occurrence might impact morphosyntactic structure and therefore, as Stage II concluded, the three factors of frequency of occurrence and co-occurrence and morphosyntactic type are interrelated.

Overall, Stage III of the research provided a comprehensive answer to Research Question 2 about how collocation-specific factors can be used for teaching and testing collocations. First, it demonstrated the importance of selecting collocational items for teaching and testing based on multiple levels: semantic (frequency of co-occurrence or the strength of association between collocational components), structural (morphosyntactic type), and usage-based (frequency of occurrence). Then, it showed the effectiveness of learning tasks and exercises that focus on the explicit analysis of collocational characteristics and reinforce noticing and repetition.

Finally, the three stages of the research provided an answer to Research Question 3 on the optimal format for testing collocations. The research confirmed that the format of the cloze and accuracy judgement-based multiple-choice matrix adapted from Revier (2009) allows to measure participants' collocational competence; first, by ensuring context; second, by asking them to identify both lexical components, that is, a whole collocation; and third, by tapping into the most common type of collocational errors of the accuracy/appropriacy of use, that is, paraphrasing (as per Schmitt, 2013).

Contributions to Collocational Studies

Although quite a few researchers have examined different factors influencing collocation acquisition, only a relatively small percentage of these studies have attempted to classify and compare these factors. While studies by Bardovi-Harlig and Bastos (2011), Carrol et al. (2016), Schmitt et al. (2004), and Yamashita and Jiang (2010) focus on external factors such as L1-L2 interference, L2 input, and proficiency, studies by Ellis et al. (2008), Nguyen and Webb (2016), Peters (2016), and Webb and Kagimoto (2011) explore collocation characteristics. Fewer studies, for example by Koya (2005), Macis and Schmitt (2017), Nesselhauf (2005), and Szudarski and Conklin (2014), combine external and internal factors. The present study follows in the steps of this final group of researchers; however, it also expands the typology of mixed factors across the three levels – extralinguistic, interlinguistic, and intralinguistic.

Another important contribution of this research is the identification and analysis of collocational characteristics that make collocations easier or more difficult to be recognized and produced by ESL learners. This study confirmed that strong frequent collocations are less challenging, while strong infrequent collocations (with higher frequency of co-occurrence) are the most challenging for ESL learners, yet this latter type is susceptible to teaching treatment. This result shows the importance of selecting collocational items for teaching both on the basis of frequency of co-occurrence and frequency of occurrence.

Additionally, to the best of my knowledge, this study is the only one that compares five morphosyntactic collocational types as related to collocation recognition and controlled production, as existing research predominantly focuses on Verb + Noun

and Adjective + Noun types only. It identified that Verb + Noun, Adjective + Noun, and Verb + Adjective collocations are the strongest score predictors, and yet Verb + Noun and Adjective + Noun collocations are also less susceptible to teaching intervention, while Verb + Adjective and Adverb + Adjective collocations are generally acquired more quickly. The study also proved the interconnectedness of the factors of morphosyntactic structure and frequency of occurrence and co-occurrence because, as Stages II and III showed, different morphosyntactic types might be recognized more or less easily depending on their normed frequency and MI scores. Thus, Adverb + Adjective and Verb + Noun collocations with low frequency of occurrence and high frequency of co-occurrence showed the most improvement in test scores from pre- to post-instruction, but in the case of Verb + Adjective collocations, it was the items with low frequency of occurrence and co-occurrence that were the most improved. This finding has not been typically addressed in the literature, and would thus benefit from more detailed analysis and further research.

This study explored different testing formats and confirmed the effectiveness of an acceptability judgement task as an adequate tool for measuring collocation recognition as per Wolter and Gyllstad (2013), Szudarski and Conklin (2014), and others. In addition, it suggested the necessity of modifying the traditional standardized format of multiple-choice tests following Revier's (2009) CONTRIX model so as to measure participants' knowledge of whole collocations in their authentic context and address collocational errors, such as paraphrasing/interchanging words with their synonyms.

Finally, this study showed that focusing on collocation characteristics while teaching and using a blended teaching approach that prioritizes repetition, noticing,

analyzing, deep processing, and active engagement as discussed in studies by Ellis (1995), Schmitt (2010), Sonbul and Schmitt (2013), Szudarski and Carter (2016), and Szudarski and Conklin (2014), leads to an improvement in participants' collocational knowledge.

Limitations and Practical Recommendations for Future Research

Most of the research limitations are related to either time constraints or the fact that this is one of very few studies that addressed a comparative typology of external and internal factors. Many of its findings have not been widely explored in the research literature; as such, this pioneering study is exploratory rather than final and conclusive. It highlights possible directions for further research on the factors influencing collocation acquisition.

One limitation is the relatively small number of factors that were covered. This is particularly noticeable at Stages II and III of the research, which target collocation characteristics. Potentially significant factors that emerged in previous studies by Durrant and Schmitt (2010), Gyllstad and Wolter (2016), Nesselhauf (2005), Nguyen and Webb (2016), Sinclair et al. (2004), and Webb and Kagimoto (2011), such as the frequency of a node and/or collocate, collocational span, degree of semantic transparency, L1 congruency, and collocations of near-synonyms, could not be addressed. In the future, these collocation characteristics should be included and correlated with the frequency of occurrence and co-occurrence and morphosyntactic structure. This would require a larger number of teaching and test items since the 18 collocations used at Stage I and the 34 collocations used at Stages II and III represented only the bare minimum and need to be expanded.

Another limitation is the lack of diversity of this study's ESL participants, most of whom who were students in the same EAP programs and consequently, had the same L1 background, age, formal instruction and amount of input in the L2, and length of residence in an English-speaking country. In order to fully explore such extralinguistic factors as L1, age, the length of residence, age of onset of learning English, and formal instruction in English, in the future it might be useful to recruit participants from different populations, proficiency levels, and from both EAP and non-academic ESL courses.

Due to the limited time frame of this study, there were no vocabulary proficiency pre-tests, and the post-test surveys at Stages II and III mostly implemented closed-ended self-assessment-based questions exploring the cognitive reading strategies that participants employed to complete the test, L1-L2 interference, and vocabulary learning strategies. Although some studies by Guduru (2014) and Kayler and Weller (2007) endorse self-assessment, it remains a somewhat subjective measurement tool. Future research might employ more complex measurements, such as eye-tracking, error rates, and reaction times for examining cognitive strategies, and Vocabulary Levels Test (Nation, 1990) or similar tests for identifying participants' proficiency levels.

Finally, since most studies (Nesselhauf, 2005; Granena & Long, 2013; Wray, 2002) point out that collocational competence develops over a long period of time, more prolonged and more intensive instruction is needed to validate the results of the teaching experiment in this study. With more time, learning tasks in future studies could transition from standardized recognition and controlled production-focused sentence-level exercises to free production-targeted writing and speaking activities and include longer reading passages to correspond to the principles of communicative and context-based language

learning (Ho & Henry, 2014; Nunan, 1991; Richards & Rogers, 2014). Additionally, in future research, the pre- and post-instruction tests should measure not only recognition and controlled production, but also free collocation production.

Overall, this research is but one step towards creating a multi-levelled comprehensive taxonomy of factors impacting collocation recognition and production by ESL learners. Further experiments would be useful to validate these results, but, when completed, such a taxonomy has the potential to revolutionize the approach to teaching collocations. Once educators and researchers know the factors of influence, they will be able to step away from the concept of the native speaker's intuition that prevails in studies about collocational competence and pinpoint more effective strategies to overcome learning challenges.

References

Abedi, Z., & Mobaraki, M. (2014). The effect of grammatical collocation instruction on understanding ESP texts for undergraduate Computer Engineering students. *Journal of Language Teaching and Research, 5*(3), 631–641.

Ackermann, K., & Chen, Y.-H. (2013). Developing the academic collocation list (ACL): A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes, 12,* 235–247.

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

Ahranjani, A. K., & Shadi, M. (2012). Input enhancement and EFL learners' collocation acquisition. *International Journal of Academic Research, 4*(1), 96-101.

Aisenstadt, E. (1979). Collocability restrictions in dictionaries. In R. R. K. Hartmann (Ed.), *Dictionaries and their users. Papers from the 1978 B.A.A.L. seminar on lexicology* (pp. 71-74). Leuven, Belgium: Katholieke Universiteit.

Akinci, M., & Yildiz, S. (2017). Effectiveness of corpus consultation in teaching verb + noun collocations to advanced ELT Students. *Eurasian Journal of Applied Linguistics, 3*(1), 91-109.

Alali, F. A., & Schmitt, N. (2012). Teaching formulaic sequences: The same as or different from teaching single words? *TESOL Journal, 3*(2), 153–180.

Alsakran, R.A. (2011). *The productive and receptive knowledge of collocations by advanced Arabic-speaking ESL/EFL learners*. (Master's Thesis). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. 1497925)

Al-Thubaity, A., & Bazeem, I. (2017). A collocation extraction tool and two language resources for MSA. *Procedia Computer Science, 117*, 23–29.

Allerton, D. J. (1984). Three (or four) levels of word co-occurrence restrictions. *Lingua, 63*, 17-40.

Almela, M. (2011). The case for verb + adjective collocations: Corpus-based analysis and lexicographical treatment. *Revista de Lingüística y Lenguas Aplicadas, 6*, 39–51.

Almela, M., Cantos, P., & Sánchez, A. (2013). Collocation, co-collocation, constellation… Any advances in distributional semantics? In Proceedings of the *5th International Conference on Corpus Linguistics* (pp. 231-240). University of Alicante, Alicante, Spain.

Attar, E. M., & Allami, H. (2013). The effects of teaching lexical collocations on speaking ability of Iranian EFL learners. *Theory and Practice in Language Studies, 3*(6), 1070-1079.

Bardovi-Harling, K., & Bastos, M. T. (2011). Proficiency, length of stay, and intensity of interaction, and the acquisition of conventional expressions in L2 pragmatics. *Intercultural Pragmatics, 8*(3), 347–384.

Barfield, A. (2009). Following individuals' L2 collocation development over time. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language. Multiple interpretations* (pp. 208-223). New York, NY: Palgrave Macmillan.

Barfield, A., & Gyllstad, H. (2009). Introduction: Researching L2 collocation knowledge and development. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language: Multiple interpretations* (pp. 1-18). New York, NY: Palgrave Macmillan.

Barnbrook, G., Mason, O., & Krishnamurthy, R. (2013). *Collocation: Applications and implications.* New York, NY: Palgrave Macmillan.

Benson, M., Benson, E., & Ilson, R. (1986). *The BBI combinatory dictionary of English.* Amsterdam, Holland: John Benjamins.

Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. In H. Hasselgard and S. Okselfjell (Eds.), *Out of corpora: Studies in honour of Stig Johansson* (pp. 181-189). Amsterdam, Holland: Rodopi.

Biber, D., Conrad, S., & Cortes, V. (2003). Lexical bundles in speech and writing: An initial taxonomy. In A. Wilson, P. Rayson, & T. McEnery (Eds.), *Corpus linguistics by the lune: A festschrift for Geoffrey Leech* (pp. 71-92). Frankfurt, Germany: Peter Lang.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English.* London, UK: Longman.

Birch, B. M. (2014). *English grammar pedagogy: A global perspective.* New York, NY: Routledge.

Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the effects of exercise on verb–noun collocations. *Language Teaching Research, 18*(1), 54–74.

Boers, F., Eyckmans, J., & Lindstromberg, S. (2013). The effect of a discrimination task on recall of L2 collocations and compounds. *International Journal of Applied Linguistics, 24*(3), 357-369.

Brandl, K. (2008). *Communicative language teaching in action: Putting principles to work.* Upper Saddle River, NJ: Pearson Prentice Hall.

Bullokar, J. (1916). *An English expositor.* London, UK: J. Legatt.

Bylund, E., Abrahamsson, N., & Hyltenstam, K. (2012). Does first language maintenance hamper nativelikeness in a second language? A study of ultimate attainment in early bilinguals. *Studies in Second Language Acquisition, 34*(2), 215-241.

Carrol, G., Conklin, K., & Gyllstad, H. (2016). Found in translation. The influence of the L1 on the reading of idioms in a L2. *Studies in Second Language Acquisition, 38*(3), 403-443.

Carter, R., & McCarthy, M. (1988). *Vocabulary and language teaching.* Abington, UK: Routledge.

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*(3), 231-251.

Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology, 14*(2), 30-49.

Cieślicka, A. (2004). Formulaic language in L2. Storage, retrieval and production of idioms by second language learners. In M. Pütz & L. Sicola (Eds.). *Cognitive processing in second language acquisition* (pp. 149-168). Philadelphia, PA: John Benjamins.

Cieślicka, A. (2006). Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research 22*(2), 115-144.

Cobb, T. (n.d.). *Web VocabProfile*. Retrieved from https://www.lextutor.ca/vp/

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin, 112*(1), 155-159.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, *29*(1), 72-89.

Cowan, L. (1989). *Towards a definition of collocation.* (Master's thesis). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. ML49083).

Cowie, A. P. (1998). Phraseological dictionaries: Some East-West comparisons. In A. P. Cowie (Ed.), *Phraseology. Theory, analysis, and applications* (pp. 209-228). Oxford, UK: Claredon Press Oxford.

Cowie, A. P., Mackin, R., & McCaig, I. R. (1983). *Oxford dictionary of English idioms.* Oxford, UK: Oxford University Press.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238.

Creswell, J. V., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research.* Thousand Oaks, CA: SAGE.

Crowther, J., Dignen, S., & Lea, D. (2002). *Oxford collocations dictionary for students of English*. Oxford, UK: Oxford University Press.

Department for Education and Skills [DfES] (2006). *Excellence and enjoyment: Learning and teaching for bilingual children in the primary years*. Colegate, Norwich: DfES Publications.

Durrant, P., & Doherty, A. (2010). Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory, 6*(2), 125-155.

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make

    use of collocations? *International Review of Applied Linguistics in Language*

    Teaching, 47(2), 157-177.

Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from

    exposure. *Second Language Research, 26*(2), 163–188.

Edele, A., Seuring, J., Kristen, C., & Stanat, P. (2015). Why bother with testing? The

    validity of immigrants' self-assessed language proficiency. *Social Science*

    Research, 52, 99–123.

Edmonds, A., & Gudmestad, A. (2014). Your participation is greatly/highly appreciated:

    Amplifier collocations in L2 English. *Canadian Modern Language Review, 70*(1),

    76-102.

Ellis, N. C. (1995). The psychology of foreign language acquisition: Implications for

    CALL. *Computer Assisted Language Learning, 8*(3), 103-128.

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(3), 375-396.

Erman, B. (2014). There is no such thing as a free combination: a usage-based study of

    specific construals in adverb–adjective combinations. *English Language and*

    Linguistics, 18(1), 109–132.

Erman, B., Denke, A., Fant, L. & Forsberg Lundell, F. (2015). Nativelike expression in

    the speech of long-residency L2 users: A study of multiword structures in L2

    English, French and Spanish. *International Journal of Applied Linguistics, 25*(2),

    160-182.

Evert, S. (2008). Corpora and collocations. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics: An International Handbook* (pp. 1212-1248). Berlin, Germany: de Gruyter Mouton.

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

Fellbaum, C. (2007). Introduction. In C. Fellbaum (Ed.), *Idioms and collocations. Corpus-based and lexicographic studies* (pp. 1-22). London, UK: Continuum.

Field, A. (2009). *Discovering statistics using SPSS*. London, UK: Sage.

Firth, J. R. (1957). Modes of meaning. In J. R. Firth **(**Ed.), *Papers in linguistics 1934 – 1951* (pp. 190-215). Oxford, UK: Oxford University Press.

Forsberg, F., & Fant, L. (2010). Idiomatically speaking: Effects of task variation on formulaic language in highly proficient users of L2 French and Spanish. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 47-70). New York, NY: Continuum International.

Forsberg Lundell, F, & Lindqvist, C. (2014). Lexical aspects of very advanced L2 French. *Canadian Modern Language Review, 70*(1), 28-49.

Foster, P. (2009). Lexical diversity and native-like selection: The bonus of studying abroad. In H. Daller, D. Malvern, P.Meara, J. Milton, B. Richards, & J. Treffers-Daller (Eds.), *Vocabulary studies in first and second language acquisition: The interface between theory and application* (pp. 91–106). Basingstoke, UK: Palgrave-Macmillan.

Foster, P., Bolibaugh, C., & Kotula, A. (2014). Knowledge of native-like selections in

    L2: The influence of exposure, memory, age of onset, and motivation in foreign

    language and immersion settings. *Studies in Second Language Acquisition, 36*(1),

    101-132.

Foster-Cohen, S. (2001). First language acquisition... second language acquisition:

    "What's Hecuba to him or he to Hecuba?" *Second Language Research, 17*(4),

    329-344.

Gablasova, D., Brezina, V., & McEnery, A. M. (2011). Collocations in corpus-based

    language learning research: Identifying, comparing, and interpreting the evidence.

    Language Learning, 67, 155-179.

Gabryś-Barker, D. (2006). Language activation in the thinking processes of a

    multilingual language user. *International Journal of Multilingualism, 3*(2), 105-

    124.

Ghasemi, Z. (2003). *Fossilization and the role of attention in adults second language

    learning.* (Master's thesis). Retrieved from ProQuest Dissertations and Theses.

    (Accession Order No. MQ89840).

Gitsaki, C. (1996). *The development of ESL collocational knowledge* (Unpublished

    doctoral dissertation). The University of Queensland, Brisbain City, Australia.

    Retrieved from http://espace.library.uq.edu.au/view/UQ:205374

Gitsaki, C. (1999). *Second language lexical acquisition: A study of the development of

    collocational knowledge*. Bethesda, MD: International Scholars.

Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research, 29*(3), 311-343.

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, UK: Oxford University Press.

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 (IRAL), 52*(3), 229-252.

Granger, S., & Menier, S. (2008). Introduction: The many faces of phraseology. In S. Granger & F. Menier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. xix-xxviii). Amsterdam, Holland: John Benjamins.

Granger, S., & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 27-49). Amsterdam, Holland: John Benjamins.

Graves, M. F. (2006). *The vocabulary book: Learning and instruction*. New York, NY: Teachers College Press.

Gravetter, F. J., & Wallnau, L. B. (2012). *Using statistics for behavioral sciences.* Belmont, CA: Wadsworth Publishing.

Greenbaum, S. (1970). *Verb-intensifier collocations in English: An experimental approach*. The Hague, Netherlands: Mouton.

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). New York, NY: Palgrave Macmillan.

Guduru, R. (2014). An overview of Saudi FFL Learners' self-assessment of vocabulary learning strategies. *Language in India, 14*(5), 242–260.

Guerrero, M. C. M. (2018). Going covert: Inner and private speech in language learning. *Language Teaching, 51*(1), 1–35.

Gyllstad, H. (2007). *Testing English collocations: Developing receptive tests for use with advanced Swedish learners.* (Unpublished doctoral dissertation). Lund University, Lund, Sweden.

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

Gyllstad, H., & Wolter, B. (2016). Collocational processing in light of the phraseological continuum model: Does semantic transparency matter? *Language Learning, 66*(2), 296-323.

Halliday, M. A. K. (1966). Lexis as a linguistic level. In C. E. Bazell, J. C. Catford, M. A. K. Halliday, & R. H. Robins (Eds.), *In memory of J. R. Firth* (pp. 148-162). London, UK: Longman.

Halliday, M.A.K., & Hasan, R. (1976). *Cohesion in English*. London, UK: Longman.

Han, Z-H. (2004). *Fossilization in adult second language acquisition*. Tonawanda, NY: Multiligual Matters.

Handl, S. (2008). Essential collocations for learners of English: The role of collocational direction and weight. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 43–66). Philadelphia, PA: John Benjamins.

Hargreaves, P. (2000). Collocation and testing. In M. Lewis (Ed.), *Teaching collocation. Further developments in the lexical approach* (pp. 205-223). Boston, MA: Thomson & Heinle.

Hasselgren, A. (1994). Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary. *International Journal of Applied Linguistics, 4*(2), 237–258.

Hausmann, F. J. (1989). Le dictionnaire de collocations. In F. J. Hausmann, H. E. Wiegand, & L. Zgusta (Eds.), *Wörterbücher, dictionaries, dictionnaires: Ein internationales Handbuch zur Lexicographie* (pp. 1010-1019). Berlin, Germany: de Gruyter Mouton.

Heng, C. S., Kashiha, H., & Tan, H. (2014). Lexical bundles: Facilitating university "talk" in group discussions. *English Language Teaching 7*(4), 1-10.

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). Eurosla Monograph Series, 2. Retrieved from http://www.eurosla.org/monographs/EM02/EM02home.php

Henriksen, B., & Stoehr, L. S. (2009). Commentary on part IV: Processes in the development of L2 collocational knowledge: A challenge for language learners, researchers and teachers. In A. Barfield & H. Gyllstad (Eds.), *Researching*

collocations in another language. Multiple interpretations (pp. 224–231). New York, NY: Palgrave Macmillan.

Herbst, T. (1996). What are collocations: Sandy beaches or false teeth? *English Studies, 77*(4), 379-393.

Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation. Further developments in the lexical approach* (pp. 47-69). Boston, MA: Thomson & Heinle.

Hill, J., Lewis, M., & Lewis, M. (2000). Classroom strategies, activities and exercises. In M. Lewis (Ed.), *Teaching collocation. Further developments in the lexical approach* (pp. 88-117). Boston, MA: Thomson & Heinle.

Ho, D. G. E., & Henry, A. (2014). A text-based approach to teaching writing in the ESL classroom. In M. B. Paradowski (Ed.), *Teaching languages off the beaten track* (pp. 209–230). Berlin, Germany: Peter Lang.

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

Hoffmann, S., & Lehmann, H. M. (2000). Collocational evidence from the British National Corpus. In J. Kirk (Ed.), *Corpora galore: Analyses and techniques in describing English* (pp. 17–32). Amsterdam, Holland: Rodopi.

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

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

Hyltenstam, K., & Abrahamsson, N. (2003). Maturational constraints in SLA. In C. J. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 539-588). London, UK: Blackwell.

Irujo, S. (1986). Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language. *TESOL Quarterly*, *20*(2), 287–304.

Jackson, M. D., & McClelland, J. L. (1979). Processing determinants of reading speed. *Journal of Experimental Psychology, 108*(2), 151-181.

Jaén, M. M. (2007). A corpus-driven design of a test for assessing the ESL collocational competence of university students. *International Journal of English Studies, 7*(2), 127-147.

JACET Committee of Basic Words Revision (2003). *JACET list of 8000 basic words.* Tokyo, Japan: The Japan Association of College English Teachers.

Jafarpour, A.A., Hashemian, M., & Alipour, S. (2013). A corpus-based approach toward teaching collocation of synonyms. *Theory and Practice in Language Studies, 3*(1), 51-60.

Jiang, J. (2009). Designing pedagogical materials to improve awareness and productive use of L2 collocations. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language. Multiple interpretations* (pp. 99-113). New York, NY: Palgrave Macmillan.

Jiang, N. A. N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal, 91*(3), 433-445.

Jones, S., & Sinclair, J. (1974). English lexical collocations. *Cahiers de Lexicologie, 24,* 15-61.

Kayler, M., & Weller, K. (2007). Pedagogy, self-assessment, and online discussion groups. *Educational Technology & Society, 10*(1), 136-147.

Kim, S. (2017). "Minimum input, maximum output, indeed!" Teaching collocations through collocation dictionary skills development. *Lexikos, 27*, 265–286.

Kjellmer, G. (1990). A mint of phrases. In K. Aijmer & B. Altenberg (Eds.), *English corpus linguistics: Studies in honour of Jan Svartvik* (pp. 111-127). London, UK: Longman.

Kjellmer, G. (1994). *A dictionary of English collocations: Based on the Brown corpus.* Oxford, UK: Claredon Press.

Koo, K. (2006). *Effects of using corpora and online reference tools on foreign language writing: a study of Korean learners of English as a second language*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. 3211793)

Korosadowicz-Struzynska, M. (1980). Word collocations in FL vocabulary instruction. *Studia Anglica Posnaniensia, 12*, 109-120.

Koya, T. (2005). *The acquisition of basic collocations by Japanese learners of English*. (Unpublished doctoral dissertation). Waseda University, Tokyo, Japan.

Koya, T. (2006). What is the reality of collocation use by native speakers of English? *Dialogue, 5*, 1-18.

Krashen, S., Long, M. H., & Scarcella, R. (1979). Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly, 13*(4), 573–582.

Krishnamurthy, R. (2006). Collocations. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (pp. 596-600). Vancouver, BC: Elsevier.

Kunin, A. V. (1984). *English-Russian phraseological dictionary*. Moscow, Russia: Russkii Yazik.

Kuo, C.-L. (2009). An analysis of the use of collocation by intermediate EFL college students in Taiwan. *ARECLS, 6*, 141-155.

Lackman, K. (2011). *Teaching collocations. Activities for vocabulary building.* Toronto, ON: Ken Lackman & Associates. Educational Consultants. Retrieved from http://www.kenlackman.com/files/Collocations11sample.pdf

Lahmann, C., Steinkrauss, R., & Schmid, M. S. (2016). Factors affecting grammatical and lexical complexity of long-term L2 speakers' oral proficiency. *Language Learning, 66*(2), 354-385.

Lai, Y.-L. (2005). *Teaching vocabulary learning strategies: awareness, beliefs, and practice. A survey of Taiwanese EFL senior high school teachers*. (Unpublished master's thesis). The University of Essex, Essex, UK.

Lakshmanan, U. (2006). Child second language acquisition and the fossilization puzzle. In Z-H. Han & T. Odlin (Eds.), *Studies of fossilization in second language acquisition* (pp. 100-133). Tonawanda, NY: Multilingual Matters Ltd.

Laufer, B., & Waldman, T. (2011). Verb-Noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning, 61*(2), 647–672.

Leech, G. N. (1974). *Semantics: The study of meaning*. Middlesex, UK: Penguin Books.

Lewis, M. (2000). Language in the lexical approach. In M. Lewis (Ed.), *Teaching collocation. Further developments in the lexical approach* (pp. 126-154). Boston, MA: Thomson & Heinle.

Li, J., & Schmitt, N. (2010). The development of collocation use academic text by
    advanced L2 learners: A multiple case study approach. In D. Wood (Ed.),
    Perspectives on formulaic language: Acquisition and communication (pp. 23-46).
    New York, NY: Continuum International.

Liao, E.-H. (2010). *An investigation in crosslinguistic transfer in EFL learners'*
    phraseology (Doctoral dissertation). Retrieved from ProQuest Dissertations and
    Theses. (Accession Order No. 3401772).

Macis, M., & Schmitt, N. (2017). Not just 'small potatoes': Knowledge of the idiomatic
    meanings of collocations. *Language Teaching Research, 21*(3), 321–340.

Mahvelati, E. H., & Mukundan, J. (2012). The effects of input flood and consciousness-
    raising approach on collocation knowledge development of language learners.
    International Journal of Applied Linguistics and English Literature, 1(6), 182-
    192.

Marks, J., & Wooder, A. (2007). *Natural English collocations*. London, UK:
    Bloomsbury.

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics, 33*(3),
    299-320.

McCarthy, M., & O'Dell, F. (2005). *English collocations in use: Advanced*. Cambridge,
    UK: Cambridge University Press.

McGee, I. (2006). *Lexical intuitions and collocation patterns in corpora*. (Unpublished
    doctoral dissertation). Cardiff University, Cardiff, UK.

Mel'čuk, I. (1998). Collocations and lexical functions. In A. P. Cowie (Ed.), *Phraseology.*
    Theory, analysis, and applications (pp. 23-54). Oxford, UK: Claredon.

Menken, K. (2013). Emergent bilingual students in secondary school: Along the academic language and literacy continuum. *Language Teaching, 46*(4), 428-476.

Metin, S. K., & Karaoğlan, B. (2011). Measuring collocation tendency of words. *Journal of Quantitative Linguistics, 18*(2), 174-187.

Mian, A. H. (1988). A microanalysis of collocation in the interlanguage of Pakistani adults learning English as a second language. (Master's thesis). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. ML47407)

Millar, N. (2011). The processing of malformed formulaic language. *Applied Linguistics, 32*(2), 129-148.

Mitchell, T. F. (1971). Linguistic 'goings-on': Collocations and other lexical matters arising on the syntactic record. *Archivum Linguisticum, 2*, 35-69.

Miyakoshi, T. (2009). *Investigating ESL learners' lexical collocations: The acquisition of verb + noun collocations by Japanese learners of English* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. 3367916)

Mochizuki, M. (1998). A vocabulary size test for Japanese learners of English. *The IRLT Bulletin, 12*, 27-53.

Moon, R. (2011). English adjectives in -like , and the interplay of collocation and morphology. *International Journal of Corpus Linguistics, 16*(4), 486-513.

Mounya, A. (2010). *Teaching lexical collocations to raise proficiency in foreign language writing* (Unpublished master's thesis). Guelma University, Guelma, Algeria.

Nation, P. (1990). *Teaching and learning vocabulary*. New York, NY: Newbury House.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge, UK:

    Cambridge University Press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening?

    Canadian Modern Language Review 63(1), 59-82.

Nation, I. S. P. (2009). Reading faster. *International Journal of English Studies, 9*(2),

    131-144.

Nation, I. S. P. (2012). *The BNC/COCA headword lists*. Retrieved from:

    https://www.victoria.ac.nz/lals/about/staff/paul-nation#vocab-tests

Nation, I. S. P., & Newton, J. (1997). Teaching vocabulary. In J. Coady & T. Huckin

    (Eds.), *Second language vocabulary acquisition* (pp. 238-254). Cambridge, UK:

    Cambridge University Press.

Nation, I. S. P., & Webb. S. A. (2011). *Researching and analyzing vocabulary*. Boston,

    MA: Heinle.

Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching.*

    Oxford, UK: Oxford University Press.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some

    implications for teaching. *Applied Linguistics, 24*(2), 223–242.

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

    Benjamins.

Nguyen, T. M. H., & Webb, S. (2016). Examining second language receptive knowledge

    of collocation and factors that affect learning. *Language Teaching Research,*

    21(3), 298–320.

Nizonkiza, D. (2012). Quantifying controlled productive knowledge of collocations across proficiency and word frequency levels. *Studies in Second Language Learning and Teaching, 2*(1), 67–92.

Nizonkiza, D. (2015). Measuring receptive collocational competence across proficiency levels. *Stellenbosch Papers in Linguistics, 44*, 125–146.

Nizonkiza, D. & Van de Poel, K. (2014). Teachability of collocations: The role of word frequency counts. *Southern African Linguistics and Applied Language Studies 32*(1), 301-316.

Nunan, D. (1991). Communicative tasks and language curriculum. *TESOL Quarterly, 25*(2), 279-295.

Office for Standards in Education, Children's Services and Skills [Ofsted] (2005, July 13). *Could they do even better? Office for Standards in Education, Children's Services and Skills Departmental Report 2008-09*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/238633/7597.pdf

Oliver, R., & Young, S. (2016). Improving reading fluency and comprehension in adult ESL learners using bottom-up and top-down vocabulary training. *Studies in Second Language Learning and Teaching, 6*(1), 111-133.

Pallant, J. (2010). *SPSS survival manual*. Berkshire, UK: Open University Press.

Palmer, H. E. (1933). *Second interim report on English collocations*. Tokyo, Japan: Kaitakusha.

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

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*(1), 75–94.

Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research, 20*(1), 113–138.

Poulsen, S. (2005). *Collocation as a language resource. A functional and cognitive study in English phraseology* (Unpublished doctoral dissertation). University of Southern Denmark, Odense, Denmark.

Radhakrishna, R. B. (2007). Tips for developing and testing questionnaires/instruments. *Journal of Extension, 45*(1). Retrieved from http://www.joe.org/joe/2007february/tt2.php

Renouf, A., & Sinclair, J. (1991). Collocational frameworks in English. In A. Karin & B. Altenberg (Eds.), *English corpus linguistics. Studies in honour of Jan Svartvik* (pp. 128-143). New York, NY: Longman.

Reppen, R. (2009). Exploring L1 and L2 writing development through collocations: A corpus based look. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language. Multiple interpretations* (pp. 49-59). New York, NY: Palgrave Macmillan.

Revier, R.L. (2009). Evaluating a new test of whole English collocations. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language* (pp. 49-59). London: Palgrave Macmillan.

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching*. Cambridge, UK: Cambridge University Press.

Rubin, G. S. (2013). Measuring reading performance. *Vision Research, 90*, 43-51.

Saldaña, J. (2012). *The coding manual for qualitative researchers.* Thousand Oaks, CA: SAGE.

Sarvari, S., Gukani, A. H., & Khomami, H. Y. (2016). Teaching collocations: Further developments in L2 speaking fluency. *Journal of Applied Linguistics and Language Research, 3*(3), 278-289.

Schmitt, D., & Schmitt, N. (2005). *Focus on vocabulary: Mastering the academic word list.* London, UK: Longman Press.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge, UK: Cambridge University Press.

Schmitt, N. (2010). *Researching vocabulary. A vocabulary research manual*. New York, NY: Palgrave McMillan.

Schmitt, N. (2013). Formulaic language and collocation. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Hoboken, NJ: Blackwell Publishing.

Schmitt, N., Dörnyei, Z., Adolphs, S., & Durow, V. (2004). Knowledge and acquisition of formulaic sequences: A longitudinal study. In N. Schmitt (Ed.), *Formulaic sequences. Acquisition, processing and use* (pp. 55-86). Philadelphia, PA: John Benjamins.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing, 18*(1), 55-88.

Shin, D., & Nation, P. (2008). Beyond single words: the most frequent collocations in spoken English. *ELT Journal, 62*(4), 339–348.

Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics, 31*(4), 487-512.

Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.

Sinclair, J. M. (2001). Collins COBUILD English language dictionary. London, UK: Collins.

Sinclair, J. M., Jones, S., & Daley, R. (2004). *English collocation studies: The OSTI report*. London, UK: Continuum.

Sitzmann, T., Ely, K., Brown, K., & Bauer, K. (2010). Self-assessment of knowledge: A cognitive learning or affective measure? *Academy of Management Learning & Education, 9*(2), 169-191.

Siyanova, A. (2010). *On-line processing of multi-word sequences in a first and second language: evidence from eye-tracking and ERP* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. U557263)

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research, 27*(2), 251-272.

Siyanova-Chanturia, A., & Martinez, R. (2014). The idiom principle revisited. *Applied Linguistics, 36*(5), 549–569.

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation:

    A multi-study perspective. *The Canadian Modern Language Review*, *64*(3), 429–

    458.

Siyanova-Chanturia, A., & Spina, S. (2015). Investigation of native speaker and second

    language learner intuition of collocation frequency. *Language Learning, 65*(3),

    533–562.

Smith, C. (2005). *The lexical approach: Collocation in high school English language*

    learners (Doctoral dissertation). Retrieved from ProQuest Dissertations and

    Theses. (Accession Order No. 3195526)

Sonbul, S. (2015). Fatal mistake, awful mistake, or extreme mistake? Frequency effects

    on off/ line/on-line collocational processing. *Bilingualism: Language and*

    Cognition, 18(3), 419–437.

Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of

    collocations under different input conditions. *Language Learning, 63*(1), 121-159.

de Souza Hodne, L. C. (2009). *Collocations and teaching: Investigating word*

    combinations in two English textbooks for Norwegian upper secondary school

    students. (Unpublished master's thesis). The University of Bergen, Bergen,

    Norway.

Stefanowitsch, A., & Gries, S. (2003). Collostructions: Investigating the interaction

    between words and constructions. *International Journal of Corpus Linguistics,*

    8(2), 209-243.

Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with

    quantitative studies. *Functions of Language, 2*(1), 23-55.

Szudarski, P., & Carter, R. (2016). The role of input flood and input enhancement in EFL

    learners' acquisition of collocations. *International Journal of Applied Linguistics,*

    26(2), 245–265.

Szudarski, P., & Conklin, K. (2014). Short- and long-term effects of rote rehearsal on

    ESL learners' processing of L2 collocations. *TESOL Quarterly, 48*(4), 833–842.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics.* Boston, MA:

    Allyn and Bacon.

Talakoob, F., & Koosha, M. (2017). Productive and receptive collocational knowledge of

    Iranian EFL learners at different proficiency levels. *International Journal of*

    Applied Linguistics and English Literature, 6(7), 11–16.

Tarone, E. (2006). Fossilization, social context and language play. In Z.-H. Han & T.

    Odlin (Eds.), *Studies of fossilization in second language acquisition* (pp. 157-

    172). Tonawanda, NY: Multilingual Matters.

Tollefson, J. W., & Firn, J. T. (1983). Fossilization in second language acquisition: An

    intermodel view. *RELC Journal, 14*(2), 19-34.

Tran, T. H. (2011). Using ready-made materials for teaching idioms. In Proceedings of

    The 41st Annual New York State TESOL Conference. Melville, NY. Retrieved

    from https://files.eric.ed.gov/fulltext/ED526846.pdf

Trochim, W.M., & Donnelly, J.P. (2006). *The research methods knowledge base.*

    Cincinnati, OH: Atomic Dog Publishing.

Vechtomova, O., Robertson, S., & Jones, S. (2003). Query expansion with long-span

    collocates. *Information Retrieval, 6*(2), 251-273.

Vigil, N. A., & Oller, J. W. (1976). Rule fossilization: A tentative model. *Language Learning, 26*(2), 281-295.

Vilkaité, L. (2016). Are non-adjacent collocations processed faster? J*ournal of Experimental Psychology: Learning, Memory, and Cognition, 42*(10), 1632-1642.

Vinogradov, V. V. (1947). Ob osnovnuikh tipakh frazeologicheskikh edinits v russom yazike. [About the basic types of phraseological units in English]. In A. A. Shakhmatov (Ed.), *Sbornik statei i materialov* [The collection of articles and materials] (pp. 339-364). Moscow, Russia: Nauka.

Vinogradov, V. V. (1977). Ob osnovnuikh tipakh frazeologicheskikh edinits v russom yazike. [About the basic types of phraseological units in English]. In V. V. Vinogradov, *Izbrannie trudi. Leksikologia i leksikografia* [Selected works. Lexicology and lexicography] (pp. 140-161). Moscow, Russia: Nauka.

Walker, C. (2008). Factors which influence the process of collocation. In F. Boers, & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and Phraseology* (pp. 291-308). Berlin, Germany: de Gruyter Mouton.

Wang, J. T., & Good. R. (2007). The repetition of collocations in EFL textbooks: A corpus study. In Proceedings of *The Sixteenth International Symposium and Book Fair on English Teaching in the Republic of China* (pp. 223-232). Taipei, China.

Wang, H. C., & Shih, S. C. (2011). The role of language for thinking and task selection in EFL learners' oral collocational production. *Foreign Language Annals*, *44*(2), 399-416.

Wanner, L., Bohnet, B., & Giereth, M. (2006). Making sense of collocations. *Computer Speech and Language, 20*(4), 609–624.

Wanner, L., Ferraro, G., & Moreno, P. (2017). Towards distributional semantics based

    classification of collocations for collocation dictionaries. *International Journal of*

    Lexicography, 30(2), 167-186.

Webb, S., & Kagimoto, E. (2011). Learning collocations: Do the number of collocates,

    position of the node word, and synonymy affect learning? *Applied Linguistics*,

    32(3), 259-276.

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language*

    Learning, 63(1), 91-120.

Weinert, R. (2010). Formulaicity and usage-based language: Linguistic, psycholinguistic

    and acquisitional manifestations. In D. Wood (Ed.), *Perspectives on formulaic*

    language. Acquisition and communication (pp. 1-22). New York, NY:

    Continuum.

West, M. (1953). *A general service list of English words*. London, UK: Longman.

de Wit, V. D. (2007). *Fossilization: A case study of an adult learner*. (Master's thesis).

    Retrieved from ProQuest Dissertations and Theses. (Accession Order No.

    0669018)

Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the

    influence of L1 intralexical knowledge. *Applied Linguistics*, 32(4), 430–449.

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing: A

    comparison of congruent and incongruent collocations. *Studies in Second*

    Language Acquisition, 35, 451-482.

Wolter, B., & Yamashita, J. (2015). Processing collocations in a second language: A case

    of first language activation? *Applied Psycholinguistics, 36*, 1193-1221.

Woolard, G. (2000). Collocation: Encouraging learner independence. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 28-46). Boston, MA: Thomson & Heinle.

Wood, D. (2010). Lexical clusters in an EAP textbook corpus. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 88-106). New York, NY: Continuum International.

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

Wray. A. (2008). *Formulaic language: Pushing the boundaries*. Oxford, UK: Oxford University Press.

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language and Communication, 20*(1), 1-28.

Xiao, R., & McEnery, T. (2006). Can completive and durative adverbials function as tests for telicity? Evidence from English and Chinese. *Corpus Linguistics and Linguistic Theory, 1*(2), 1-21.

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*(4), 647–668.

Yan, H. (2010). Study on the causes and countermeasures of the lexical collocation mistakes in college English. *English Language Teaching, 3*, 162–165.

Yin, R. (2009). *Case study research: Design and methods*. Thousand Oaks, CA: Sage.

Ying, Y., & O'Neill, M. (2009). Collocation learning through an 'AWARE' approach: Learner perspectives and learning process. In A. Barfield & H. Gyllstad (Eds.),

Researching collocations in another language. Multiple interpretations (pp. 181-193). New York, NY: Palgrave Macmillan.

Yorio, C. A. (1989). Idiomaticity as an indicator of second language proficiency. In K. Hyltenstam, & L. K. Obler (Eds.), *Bilingualism across the lifespan* (pp. 55-72). Cambridge, UK: Cambridge University Press.

Xu, J. (2015). Predicting ESL learners' oral proficiency by measuring the collocations in their spontaneous speech. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. 10009317).

Zhang, X. (2017). Effects of receptive-productive integration tasks and prior knowledge of component words on L2 collocation development. *System, 66*, 156–167.

Zhu, F., & Fellbaum, C. (2015). Quantifying fixedness and compositionality in Chinese idioms. *International Journal of Lexicography, 28*(3), 338–350.

Ziefle, M. (1998). Effects of display resolution on visual performance. *Human Factors, 40*(4), 554-568.

Appendices

Appendix A: Ethics Clearance Forms

Carleton University Research Office
Research Ethics Board
1325 Dunton Tower
1125 Colonel By Drive
Ottawa, ON K1S 5B6 Canada
Tel: 613-520-2517
ethics@carleton.ca

Ethics Clearance Form

This is to certify that the Carleton University Research Ethics Board has examined the application for ethical clearance. The REB found the research project to meet appropriate ethical standards as outlined in the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, 2nd edition* and, the *Carleton University Policies and Procedures for the Ethical Conduct of Research*.

X New clearance
☐ **Renewal of original clearance**   Original date of clearance:

Date of clearance	**4 December 2013**
Researcher	**Olga Makinina, Ph.D. student**
Department	**School of Linguistics and Language Studies**
Supervisor	**Prof. Janna Fox and Prof. Natasha Artemeva, School of Linguistics and Language Studies**
Project number	**100692**
Title of project	**Collocation recognition by adult speakers of English as a first and second language**

Clearance expires: **31 May 2014**

All researchers are governed by the following conditions:

Annual Status Report: You are required to submit an Annual Status Report to either renew clearance or close the file. Failure to submit the Annual Status Report will result in the immediate suspension of the project. Funded projects will have accounts suspended until the report is submitted and approved.

Changes to the project: Any changes to the project must be submitted to the Carleton University Research Ethics Board for approval. All changes must be approved prior to the continuance of the research.

Adverse events: Should any participant suffer adversely from their participation in the project you are required to report the matter to the Carleton University Research Ethics Board. You must submit a written record of the event and indicate what steps you have taken to resolve the situation.

Suspension or termination of clearance: Failure to conduct the research in accordance with the principles of the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, 2nd edition* and the *Carleton University Policies and Procedures for the Ethical Conduct of Research* may result in the suspension or termination of the research project.

Andy Adler, Chair
Carleton University Research Ethics Board
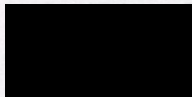
Louise Heslop, Vice-Chair
Carleton University Research Ethics Board

Carleton University Research Office
Research Ethics Board
1325 Dunton Tower
1125 Colonel By Drive
Ottawa, ON K1S 5B6 Canada
Tel: 613-520-2517
ethics@carleton.ca

Ethics Clearance Form – Clearance Renewal

This is to certify that the Carleton University Research Ethics Board has examined the application for ethical clearance. The REB found the research project to meet appropriate ethical standards as outlined in the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Human, 2nd edition*, and the *Carleton University Policies and Procedures for the Ethical Conduct of Research*.

Original Date of Clearance: 4 December 2013
Renewal Date of Clearance: May 02, 2014
Researcher: Olga Makinina (Student - Phd)
Department: Faculty of Arts and Social Sciences\Linguistics and Applied Language Studies (School of)
University: Carleton University
Research Supervisors (if applicable): Janna Dorothy Fox and Natalia Artemeva
Project Number: 100692
Alternate File Number (if applicable):
Project Title: Recognition of false collocations by speakers of English as a first and additional language

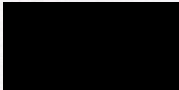**Clearance Expires:** May 31, 2015

_____

All researchers are governed by the following conditions:

Annual Status Report: You are required to submit an Annual Status Report to either renew clearance or close the file. Failure to submit the Annual Status Report will result in the immediate suspension of the project. Funded projects will have accounts suspended until the report is submitted and approved.

Changes to the project: Any changes to the project must be submitted to the Carleton University Research Ethics Board for approval. All changes must be approved prior to the continuance of the research.

Adverse events: Should a participant suffer adversely from their participation in the project you are required to report the matter to the Carleton University Research Ethics Board. You must submit a written record of the event and indicate what steps you have taken to resolve the situation.

Suspension or termination of clearance: Failure to conduct the research in accordance with the principles of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, 2nd edition and the Carleton University Policies and Procedures for the Ethical Conduct of Research may result in the suspension or termination of the research project.

313

Andy Adler
Chair, Carleton University Research Ethics Board

Louise Heslop
Vice-Chair, Carleton University Research Ethics Board

Carleton University
Research Ethics Office
Research Ethics Board
511 Tory, 1125 Colonel By Drive
Ottawa, ON K1S 5B6 Canada
Tel: 613-520-2517, ethics@carleton.ca

Ethics Clearance Form – Clearance Renewal

This is to certify that the Carleton University Research Ethics Board has examined the application for ethical clearance. The REB found the research project to meet appropriate ethical standards as outlined in the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Human, 2nd edition*, and the *Carleton University Policies and Procedures for the Ethical Conduct of Research*.

Original Date of Clearance: December 4, 2013
Renewal Date of Clearance: May 01, 2015
Researcher: Olga Makinina (Student Research: Ph.D. Student)
Department: Faculty of Arts and Social Sciences\Linguistics and Applied Language Studies (School of)
University: Carleton University
Research Supervisor (if applicable): Janna Dorothy Fox and Natalia Artemeva (Research Supervisor)
Project Number: 100692
Alternate File Number (if applicable):
Project Title: Recognition of false collocations by speakers of English as a first and additional language
Funder (if applicable):

Clearance Expires: May 31, 2016

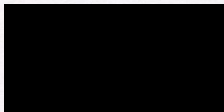**All researchers are governed by the following conditions:**

Annual Status Report: You are required to submit an Annual Status Report to either renew clearance or close the file. Failure to submit the Annual Status Report will result in the immediate suspension of the project. Funded projects will have accounts suspended until the report is submitted and approved.

Changes to the project: Any changes to the project must be submitted to the Carleton University Research Ethics Board for approval. All changes must be approved prior to the continuance of the research.

Adverse events: Should a participant suffer adversely from their participation in the project you are required to report the matter to the Carleton University Research Ethics Board. You must submit a written record of the event and indicate what steps you have taken to resolve the situation.

Suspension or termination of clearance: Failure to conduct the research in accordance with the principles of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, 2nd edition and the Carleton University Policies and Procedures for the Ethical Conduct of Research may result in the suspension or termination of the research project.

Louise Heslop
Chair, Carleton University Research Ethics Board
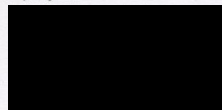
Andy Adler
Vice-Chair, Carleton University Research Ethics Board

Research Compliance Office
511 Tory | 1125 Colonel By Drive
| Ottawa, Ontario K1S 5B6
613-520-2600 Ext: 2517
ethics@carleton.ca

Dear Olga Makinina,

Thank you for your Change to Protocol request submitted to CUREB-A on July 12, 2016 for your ethics protocol entitled **Recognition of false collocations by speakers of English as a first and additional language [Olga Makinina] (Project #100692).** The changes you are proposing are acceptable and you may proceed.

Effective: **July 13, 2016**   Expires: **June 30, 2017**.

Restrictions:

This certification is subject to the following conditions:

1. Approval is granted only for the research and purposes described in the application.

2. Any modification to the approved research must be submitted to CUREB-A. All changes must be approved prior to the continuance of the research.

3. An Annual Application for the renewal of ethics clearance must be submitted and approved by the above date. Failure to submit the Annual Status Report will result in the closure of the file. If funding is associated, funds will be frozen.

4. A closure request must be sent to CUREB-A when the research is complete or terminated.

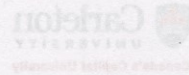**5.** Should any participant suffer adversely from their participation in the project you are required to report the matter to CUREB-A.

6. Failure to conduct the research in accordance with the principles of the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans 2nd edition* and the *Carleton University Policies and Procedures for the Ethical Conduct of Research* may result in the suspension or termination of the research project.

Please email the Ethics Coordinators at ethics@carleton.ca if you have any questions.

APPROVED BY: **Date: July 13, 2016**

Andy Adler, PhD, Chair, CUREB-A
Shelley Brown, PhD, Vice Chair, CUREB-A

Research Compliance Office
511 Tory | 1125 Colonel By Drive
| Ottawa, Ontario K1S 5B6
613-520-2600 Ext: 4085
ethics@carleton.ca

CERTIFICATION OF INSTITUTIONAL ETHICS CLEARANCE

A Change to Protocol for the following research has been **cleared** by the Carleton University Research Ethics Board-A (CUREB-A) at Carleton University. The researcher may proceed with their research. CUREB-A is constituted and operates in compliance with the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* (TCPS2).

Ethics Protocol Clearance ID: Project # 100692
Faculty Supervisor: David Wood, Carleton University
Research Team (and roles) (If applicable): **David Wood (Primary Investigator);** Olga Makinina (Student Researcher); Peggy Hartwick (Collaborator)

Study Title: Recognition of false collocations by speakers of English as a first and additional language [Olga Makinina]
Funding Source (If applicable):
Effective: **November 11, 2016**   Expires: **June 30, 2017**.

Please email the Ethics Coordinators at ethics@carleton.ca if you have any questions or if you require a copy with a signature.

CLEARED BY:	**Date:**
Andy Adler, PhD, Chair, CUREB-A	**November 11, 2016**

Shelley Brown, PhD, Vice-Chair, CUREB-A

Research Compliance Office
511 Tory | 1125 Colonel By Drive
| Ottawa, Ontario K1S 5B6
613-520-2600 Ext: 4085
ethics@carleton.ca

CERTIFICATION OF INSTITUTIONAL ETHICS CLEARANCE

The Carleton University Research Ethics Board-A (CUREB-A) at Carleton University has renewed ethics approval for the research project detailed below. CUREB-A is constituted and operates in compliance with the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* (TCPS2).

Title: Recognition of false collocations by speakers of English as a first and additional language [Olga Makinina]

Protocol #: 100692

Project Team Members: David Wood (Primary Investigator)
Olga Makinina (Student Researcher)
Peggy Hartwick (Collaborator)

Department and Institution: Faculty of Arts and Social Sciences\Linguistics and Language Studies (School of), Carleton University

Funding Source (If applicable):

Effective: **May 26, 2017**   Expires: **May 31, 2018**.

Restrictions:

This certification is subject to the following conditions:

1. Clearance is granted only for the research and purposes described in the application.
2. Any modification to the approved research must be submitted to CUREB-A. All changes must be approved prior to the continuance of the research.
3. An Annual Application for the renewal of ethics clearance must be submitted and cleared by the above date. Failure to submit the Annual Status Report will result in the closure of the file. If funding is associated, funds will be frozen.
4. A closure request must be sent to CUREB-A when the research is complete or terminated.
5. Should any participant suffer adversely from their participation in the project you are required to report the matter to CUREB-A.
6. It is the responsibility of the student to notify their supervisor of any adverse events, changes to their application, or requests to renew/close the protocol.

7. Failure to conduct the research in accordance with the principles of the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans 2nd edition* and the *Carleton University Policies and Procedures for the Ethical Conduct of Research* may result in the suspension or termination of the research project.

Please email the Research Compliance Coordinators at ethics@carleton.ca if you have any questions or if you require a clearance certificate with a signature.

CLEARED BY: **Date: May 26, 2017**

Andy Adler, PhD, Chair, CUREB-A
Shelley Brown, PhD, Vice-Chair, CUREB-A

Research Compliance Office
511 Tory | 1125 Colonel By Drive
| Ottawa, Ontario K1S 5B6
613-520-2600 Ext: 4085
ethics@carleton.ca

CERTIFICATION OF INSTITUTIONAL ETHICS CLEARANCE

The Carleton University Research Ethics Board-A (CUREB-A) has granted ethics clearance for changes to protocol to the research project described below and research may now proceed.

CUREB-A is constituted and operates in compliance with the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* (TCPS2).

Ethics Clearance ID: Project # 100692

Project Team Members: **David Wood (Primary Investigator)**
Olga Makinina (Student Researcher)
Peggy Hartwick (Collaborator)

Project Title: **Factors influencing collocation acquisition by speakers of English as a first and second language [Olga Makinina]**

Funding Source (if applicable):

Effective: **April 04, 2018**   Expires: **May 31, 2018**

Upon reasonable request, it is the policy of CUREB, for cleared protocols, to release the name of the PI, the title of the project, and the date of clearance and any renewal(s).

Please email the Research Compliance Coordinators at ethics@carleton.ca if you have any questions or if you require a clearance certificate with a signature.

CLEARED BY: **Date: April 04, 2018**
Andy Adler, PhD, Chair, CUREB-A
Bernadette Campbell, PhD, Vice Chair, CUREB-A

Appendix B: Written Consent Forms

Written Consent Form. Stage I: Pilot Study

Carleton
U N I V E R S I T Y
Canada's Capital University

Title: Word Processing

Date of ethics clearance: December 4, 2013

Ethics Clearance for the Collection of Data Expires: May 31, 2014

I _____, choose to participate in a study on vocabulary understanding in text. **The researcher for this study is Olga Makinina in the Carleton University School of Linguistics and Language Studies.** She is working under supervision of Natasha Artemeva and Janna Fox, professors in the Carleton University School of Linguistics and Language Studies.

This study involves an approximately 10-20-minute exercise and a 45-minute one-on-one interview. In the exercise, you will be asked to read a text and circle/underline word combinations that seem odd/non-English-like for you. Then, you will be asked to read 10 sentences with one word per sentence missing. You will need to fill in the blanks in each sentence with one of the three words provided below. The interview will contain questions about this exercise, as well as general questions about language(s) that you know and your education level.

With your permission, the interview will be audio-recorded. I need this audio-recording for the purpose of retaining the necessary information in details. No later than a month after our interview, I will transcribe the contents of the audio-recording, and then destroy the recording.

All the information that I obtain from the exercise and your interview will be kept anonymous. All identifying information will be removed from transcripts and test, and they will be coded numerically. Research data will only be accessible by me and my supervisors. Your name and any other personal information will not be disclosed to my supervisors. I will not use your name or identifying information in any presentations or published findings of my research. The hard-copy data will be stored in a locked cabinet in a safe location. The electronic data will be stored on a personal password-protected USB Drive under a code name.

321

Once the project is completed, all research data will be kept for five years and I may use this data for other research projects on this same topic. At the end of five years, all research data will be destroyed (Electronic data will be erased and hard copies will be shredded.)

After the exercise and interview, I will be happy to provide a debriefing about my research upon your request. I hope that this research will benefit other doctoral students and professors by increasing our understanding of strategies implemented in reading and understanding vocabulary.

Your participation in this research is voluntary. You are free to withdraw from participation in my study at any time and for any reason but no later than one week after the scheduled test and interview. If you decide to withdraw, you need to give me a notice via email.

This project was reviewed by the Carleton University Research Ethics Board, which provided clearance to carry out the research. Should you have questions or concerns related to your involvement in this research, please contact:

REB contact information:
Professor Andy Adler, Chair
Professor Louise Heslop, Vice-Chair
Research Ethics Board
Carleton University
1325 Dunton Tower
1125 Colonel By Drive
Ottawa, ON K1S 5B6
Tel: 613-520-2517
ethics@carleton.ca

Researcher contact information	**Supervisor contact information:**
Name: Olga Makinina	Name: Natasha Artemeva, Janna Fox
Department: School of Linguistics and Language Studies	Department: School of Linguistics and Language Studies
	Carleton University
Carleton University	Tel: 613-520-2600 x 7452; 613-520-2600 x 2046
Tel: …………………………	Email: natasha_artemeva@carleton.ca; janna_fox@carleton.ca
Email: …………………………	

_____   _____

Signature of participant   Date

_____   _____

Signature of researcher   Date

Email Consent Form. Stage I

Dear Professor _____,

My name is Olga Makinina. I am a doctoral student in the Applied Linguistics and Discourse Studies program at Carleton University. I will be conducting a research study for my doctoral dissertation. This research, "Collocation Recognition by Adult Speakers of English as a First and Second Language", examines how proficiency level in English, primary language background and text characteristics impact collocation recognition. Collocation is a fixed word combination with a specific meaning that can be partially derived from its components. For my research, I will need to administer an approximately 10-minute test and a 15-minute survey to your students.

I would like ask for your permission to administer the test and survey in your class at the time and date convenient for you. For the purpose of the test validity, it is important that students do not know in advance what they will be tested for. However, immediately following the test and survey, I am willing to give a short presentation about my doctoral studies and research in collocation acquisition.

My project was reviewed by the Carleton University Research Ethics Board, which provided clearance to carry out the research. Should you have questions or concerns related to your students' involvement in this research, please contact Andy Adler at ethics@carleton.ca or at 613-520-2517.

I attached the description and methodology/procedures of my research. Please contact me if you need any additional information at _____.

Thank you for your time and participation.

Email Consent Form. Stage II

Dear Professor _____,

I would like to ask for your permission to conduct a research study in your EAP class. The doctoral research project "Characteristics of Collocation and its Acquisition by Speakers of English as a Second/Additional Language" examines how lexical (language) characteristics of academic English collocations influence their acquisition (recognition and production). An approximately 30-40-minutes-long collocation-recognition-based test and a 10-15-minutes-long post-test survey will be administered. The test has a multiple-choice/cloze format and focuses on frequent academic English collocations. Collocations selected for the test are widely used across different academic texts and disciplines and, therefore, could have an academic value and expand the participants' vocabulary. The results of the test will be analyzed to inform the research on what collocations are the most/least challenging for participants, and what collocation characteristics might influence its acquisition. Following the test, participants will complete a survey that focuses on collocations that seemed most/least challenging for them. My project was reviewed by the Carleton University Research Ethics Board, which provided clearance to carry out the research.

Should you have questions or concerns related to your students' involvement in this research, please contact Andy Adler at ethics@carleton.ca or at 613-520-2517.

The copies of the test and survey are attached here. Please contact me if you need any additional information at _____.

Thank you for your time and participation.

Written Consent Form. Stage III

Title: Vocabulary acquisition

Date of ethics clearance: To be determined by the REB (original project cleared on December 4, 2013; Change to protocol cleared on November 11, 2016)

Ethics Clearance for the Collection of Data Expires: June 30th, 2017

I _____ (TYPE YOUR FULL FIRST AND LAST NAME), choose to participate in a study on vocabulary acquisition. **The researcher for this study is Olga Makinina in the Carleton University School of Linguistics and Language Studies.** She is working under supervision of David Wood, professor in the Carleton University School of Linguistics and Language Studies.

This study involves using results of two non-graded vocabulary skills diagnostic tests on collocations.

Your participation in this study is completely voluntary, and your refusal to participate will not affect your grade in this course. Your instructor will not know if you chose to participate in the study or not until the final course grade is made available to you and the period of the formal appeal of grade is over.

Precautions will be taken to protect your identity. All identifying information will be removed from the two skills diagnostic tests, and they will be coded numerically. Research data will only be accessible by the researcher and her supervisor. Your name and any other personal information will not be disclosed to anyone but the researcher. Your name or identifying information will not be used in any presentations or published findings of the research. The hard-copy data will be stored in a locked cabinet at Carleton University. The electronic data will be stored on a personal password-protected USB Drive under a code name.

Once the project is completed, all research data will be kept for five years and the researcher may use this data for other research projects on this same topic. At the end of five years, all research data will be destroyed (Electronic data will be erased and hard copies will be shredded.)

The ethics protocol for this project was reviewed by the Carleton University Research

Ethics Board, which provided clearance to carry out the research. If you have any ethical concerns with the study, please contact Dr. Andy Adler, Chair, Carleton University Research Ethics Board-A (by phone at 613-520-2600 ext. 2517 or via email at ethics@carleton.ca).

Researcher contact information
Name: Olga Makinina
Department: School of Linguistics and Language Studies
Carleton University
Email: …………………………..

Supervisor contact information:
Name: David Wood
Department: School of Linguistics and Language Studies
Carleton University
Tel: 613-520-2600 x 6684
Email: david_wood@carleton.ca

_____   _____

Signature of participant   Date

_____   _____

Signature of researcher   Date

Appendix C: Acceptability Judgment Task

Instructions: Please read the text below and underline the 2-to-3-words long combinations that seem "odd," not English-like, or "unnatural-sounding" to you. Please do not underline one word only or a whole sentence. You will be timed. If you finished reading before the assigned time, please do not reread the text again. Read the text only once at your normal speed.

***

Three children were in critical condition in hospital and six others were in serious but <u>balanced condition</u> after the school bus crashed into a truck north of Calgary, Alberta Health Services said. The children <u>fluctuated in age</u> from 5 to 13. "My heart goes out to the students and parents directly impacted by this incident. We are very concerned about the <u>good-being</u> of our school communities," said Susan William, acting superintendent of the Rocky View School Division. "We're not quite sure what the <u>supplying factor</u> to this collision was, whether that's <u>finite visibility</u> or road conditions or what's actually there," the police officers said. (Graveland. 2013)

***

At least 35 people in several units of a Hurdman-area rowhouse are homeless after a fast moving fire that broke out Friday morning. Eight rowhouse units at 211 Lees Ave. were damaged at a <u>guessed cost</u> of $500,000. When firefighters arrived, the fire had spread to three of the 14 units, and several vehicles parked near the building had also <u>captured fire</u>. One firefighter was taken to the hospital for an <u>unimportant injury</u>, and no one else was hurt. The cause of the fire is not yet known. (CBC News, September 27, 2013)

***

Environment Canada says thick rain and possibly thunderstorms are on the way Friday and Saturday for western areas of southern Manitoba. The agency has issued a special weather statement for: A storm system over the southwestern U.S. will hit southern Manitoba Friday night. Rain is expected over the southwest part of the province Friday morning and continuing into Friday night, with rainfall amounts of 40 to 60 mm in the southwest Manitoba and taller amounts possible locally as well. Acute thunderstorms are also possible across southern Manitoba Friday afternoon. (CBC News, October 10, 2013)

***

Canadian Pacific says there is no indication that a train derailment east of Sudbury on Sunday produces a danger to the public or the environment. Images of the scene showed the bridge collapsed and a number of cars carrying containers fell into the Wahnapitae River. Canadian Pacific spokesperson Ed Greenberg says introductory inspection of the containers indicates there are no materials or products that are dangerous. He says Canadian Pacific crews are working with native officials to determine the condition of the containers. The beginning investigation has found one of the rail cars had an unexpected wheel bearing failure that caused the derailment just before the bridge, Greenberg said in an email. (CBC News, June 2, 2013)

***

Quebec students have been in the streets this spring protesting against tuition increase, which would promote fees 75 per cent more over five years but still leave the province with rates lower than the Canadian average. While many students have returned to their classrooms, the proceeding protests have disrupted subway service in Montreal. There were warnings Thursday that the conflict is getting costly — aside from the potential

price tag of <u>enlarging the semester.</u> Provincial police said the protests have already cost them $1.5 million just in overtime. (Hughes, 2012)

References

CBC News (June 2, 2013). CP says no risk to public after several cars topple into Wahnapitae River. *CBC News*. Retrieved from http://www.cbc.ca/news/canada/sudbury/train-derailsin-bridge-collapse-near-sudbury-ont-1.1301044

CBC News (September 27, 2013). At least 35 homeless after fast-moving rowhouse fire. *CBC News*. Retrieved from http://www.cbc.ca/news/canada/ottawa/at-least-35-homeless-afterfast-moving-rowhouse-fire-1.1870403

CBC News (October 10. 2013). Southern Manitoba to get heavy rain, thunderstorms. *CBC News*. Retrieved from http://ca.news.yahoo.com/southern-manitoba-heavy-rain-thunderstorms171738166.html

Graveland, B. (October 25, 2013). Three children in critical condition, several others hurt, in bus crash near Calgary. *The Canadian Press*. Retrieved from http://kitchener.ctvnews.ca/three-children-in-critical-condition-several-others-hurt-inbus-crash-near-calgary-1.1513301

Hughes, G. (April 19, 2012). 150 arrested during escalating Quebec student protest. *The Globe and Mail*. Retrieved from https://beta.theglobeandmail.com/news/national/150-arrestedduring-escalating-quebec-studentprotest/article4101266/?ref=http://www.theglobeandmail.com&

Appendix D: List of Test Items and Their False Counterparts

Test items	*False counterparts*
Stable condition	Balanced condition
Ranged in age	Fluctuated in age
Well-being	Good-being
Contributing factor	Supplying factor
Limited visibility	Finite visibility
Estimated cost	Guessed cost
Catch fire	Capture fire
Minor injury	Unimportant injury
Heavy rain	Powerful rain
High amount	Tall amount
Severe thunderstorm	Acute thunderstorm
Pose a danger	Produce a danger
Preliminary inspection	Introductory inspection
Local official	Native official
Initial investigation	Beginning investigation
Raise fees	Promote fees
Ongoing protest	Proceeding protest
Extend the semester	Enlarge the semester

Appendix E: Cloze and Multiple Choice Task

1. The politician's speech accepting the leadership of the party took a while to get going, and his biggest applause line - hey, this is Canada - came when he said, "I will _____ the promise of universal health care."

a. preserve

b. keep

c. conserve

2. Anthony is one of those rare singers whose _____ smoking has enhanced his gifts.

a. hard

b. harsh

c. heavy

3. What matters most to mothers returning to work? Getting time off, getting _____ schedules.

a. adaptable

b. flexible

c. compliant

4. _____ drug reactions take more than 100,000 lives and add $136 billion to U.S. health-care costs each year.

a. adverse

b. conflicting

c. antagonistic

5. They must have jobs and income. Immigrants account for two-thirds of our population _____.

a. rise

b. advance

c. growth

6. Advertising will change significantly over the next couple of decades, although there's a good chance you won't notice the difference, since the most meaningful changes won't be visible to the _____ observer.

a. incidental

b. casual

c. irregular

7. The key sources of the new information are John F. and Michael K. Both were in a position to have _____ knowledge of what happened.

a. spacious

b. wide

c. extensive

8. There is the plastic _____ and hair replacement the 69-year-old billionaire has undergone to help mask the physical toll of his job.

a. operation

b. surgery

c. procedure

9. On April 18, the Seattle Times ran a photo _____ by an employee of a defense contractor in Kuwait of a plane filled with coffins.

a. taken

b. made

c. produced

10. I am constantly telling families that the real trick to not getting sick is eating healthy, getting vaccinated, and exercising. If we can get more kids and their families to _____ this advice, the next generation of adults will be healthier.

a. pursue

b. track

c. follow

References

Davies, M. *TIME magazine corpus.* Birgham Young University. Retrieved from

http://corpus.byu.edu/time/

Appendix F: Sample Interview Questions and Transcript Excerpt

Sample Interview Questions

Part 1. Immediate Discussion of the Test

What was challenging/easy for you in the test?

Was the time sufficient to finish reading and comprehend information in the text?

Did you read at your normal speed, slower, or faster than usual?

What did you focus your attention on while you were identifying incorrect word combinations? Guiding questions: Did you pay more attention to the meaning or form of the words? Did you look at each word individually, word combinations, or whole sentences/paragraphs?

Show the word combinations that you were not sure about/wanted to underline but did not. What stopped you?

The rest of the first part of the interview focuses on reading the test once again line-by-line and discussing each of participants' choices in more detail. The focus is on the "why" part: Why did you feel that this word combination is incorrect ("not-English-like")?

Part 2. Demographics and Language Background

How old are you?

What is your academic program?

Have you previously attended another university/got a degree?

Is English your first language?

If yes:

- Do you speak any other languages?

- Do you speak these languages as well as you speak English?

- What is the language you speak best after English?

If no:

- What is your first language? Is your English proficiency as good as proficiency in your first language?

- Do you speak English at home?

- At what age and where did you start learning English?

- Did you have your primary/secondary/higher (if applicable) education in English?

- Why did you start learning English? Was it your choice or school requirement?

Part 3. English Language Proficiency and Skills (for ESL speakers only)

How would you rate your knowledge of English on the scale from 1 (*poor*) to 5 (*excellent*)? Why?

How would you rate your reading/writing/listening-comprehension/speaking skills in English on the scale from 1 (*poor*) to 5 (*excellent*)? Why?

What is the most challenging for you when reading/speaking/writing/listening in English?

Does your first language influence your speaking/writing/reading/listening-comprehension in English?

If you often switch between languages, in what cases/situations do you predominantly use English, and in what cases/situations do you use your first language?

Part 4. Vocabulary Learning Strategies (for ESL speakers only)

When you are learning English words, how is it easier for you to memorize them? Individually? In phrases? In connection with other words (Which words - thematically/grammatically similar, or by association?)

When you are reading/speaking/writing/listening in English, do you focus on each word individually or word combinations?

What dictionary do you use more often: English-English or English-Your first language?

In which cases do you use English-English dictionary?

Sample Interview Transcript Excerpt[2]

O: Were the reading task and multiple-choice exercise difficult?

Interviewee Z: Sometimes I had problems because most of the times I just guessed what could be wrong, and when I saw a sentence I just covered, and every sentence a mistake. I don't know why but when I start thinking about the words and if they fit well or not, every single word made sense, and the gaps are here as well, so yeah...It was kind of difficult.

O: So, would you say you took a guess?

Interviewee Z: Yeah, it was more guessing. It was more guessing because I think I don't have English level where I can distinguish what's right and what's wrong, you know. It was all...yeah...most of the...yeah...it was just guessing.

O: You mean, for example, if you just read this text in the newspaper you wouldn't really think about this?

Interviewee Z: No, no, I wouldn't think about it, and I... I had no clue that could be a mistake or something like that.

O: Okay, so was the task difficult?

Interviewee Z: For me it was difficult, it was pretty difficult because normally I don't... when I read text or newspapers, I am not looking for mistakes because I assume that it's all written in a very good English, in a very good English, so I don't think about mistakes.

O: So if you were to rate the task from one to five, and one would be the least difficult, and five would be the most difficult, then, how would you rate it?

Interviewee Z: Um... because I don't think I have discovered all mistakes. I am pretty sure that I just have discovered like probably 10 percent or 20 percent, I think it was like four.

O: Four?

---

[2] Participants' original sentence structure, wording, and grammatical patterns were preserved.

Interviewee Z: Yeah.

O: Okay, but in that case, what was easier for you, reading the text or completing the multiple-choice exercise?

Interviewee Z: I think the exercise was easier.

O: Interesting... Why?

Interviewee Z: Um... That's a good question. I'm a little bit unsure, but I think because, um, why was it easier? Because yeah, in the text I had more possibilities, you know, it could be every single word, and here I had only three options, and yeah, so yeah, it was easier. I had only three options, so here I have thought that every single word could be wrong.

O: Yeah, that can be difficult for sure. So, if you were to compare these five paragraphs in the first task, which was the most difficult for you?

Interviewee Z: The most difficult one was the first one because there were lots of indirect speech, and I don't know... is it a slang, is it just a slang, or can you use that in this context? Is it language that you normally don't write? Is it just this... can you use these words in speech, but not in writing?

O: Oh, I see. What about other paragraphs? Were they easier?

Interviewee Z: Yeah, the other one were easier, eh... I think, what text was the text with the students?

O: About the students? The last one?

Interviewee Z: I think that was pretty easy. Um... But I haven't discovered lots of mistakes there. It was like history, it was very well written, and it was very fluent to read, so there are not lots of interruptions, like in the first text where there's indirect speech and yeah... So, um...yeah...it was pretty...um...most easy to understand. Ok, it was easy to understand but not find the mistakes.

O: I was wondering if the time was sufficient for you? You had about 20 minutes.

Interviewee Z: Yeah, the time and... There was enough time to do it. Yeah, I think the time was enough, I don't have problems.

O: So, were you reading at your normal speed?

Interviewee Z: Yeah.

O: If you had been given more time, do you think that would have changed your answers somehow?

Interviewee Z: Yeah, because I would have... I would have more time to think about everything in sentence, especially here, while reading texts. I had more time to think about it. But I don't know, I would underline some words that I've read just because I would think about it.

O: When you were reading the text, what were you focusing on when you were selecting those non-English-like word combinations?

Interviewee Z: Yeah, well, um... sometimes I was a little bit unsure but just if this sounds just awkward to me because I don't know that word or because I was right. It was always like can you say this in this context because I don't know this word, so is it right here or is it wrong?

O: So, is it because you were not really sure...

Interviewee Z: Probably here. I don't know that word [*fluctuated*]. I have never heard of it in this context. Should be children range age from 5 to 13. This is what I have like in my head and in my mind, but children fluctuated in age from 5 to 13? I've never heard it before, and I don't know that word. I can't ... I know what it means like because of the context but because I don't know that word, yeah. And here, probably I know, yeah, well-being [*good-being*] but I've never heard before *good-being*.

O: So, you were focusing on what you've heard before?

Interviewee Z: Yeah, what I've heard, yeah.

O: Did you ... did you look at words individually, word-by-word, or did you read in combinations of two-three words, or did you read the whole sentence?

Interviewee Z: I... I read the whole sentence, and then I... yeah, the fluency of the whole sentence, and then I decide is it right or wrong.

O: And then, after that, you made a decision what was right or wrong about this sentence?

Interviewee Z: Yeah. Because this text was probably easier because the sentence very short, you know, and here a little bit longer, and indirect speech, and that make it...

O: More difficult?

Interviewee Z: Yeah!

O: When you were selecting non-English-like word combinations, did you pay more attention to their meaning, or did you focus on word form and grammar, or maybe, something else?

Interviewee Z: I looked first if I have heard this before, in this context, and second, I tried to translate it. Like, I know what I mean in my language, but in this context for me it doesn't make sense.

O: So, you mean you translated the words?

Interviewee Z: Yeah, in my mind, I translated it to my mother tongue. In my mind, yeah.

O: So, you didn't focus on grammar and word form at all?

Interviewee Z: No, I didn't thought about the grammar. That would help me probably to find the mistakes to think about the grammar.

O: Are there any word combinations that you wanted to select but did not because you were unsure about them?

Interviewee Z: I was unsure about of "that incident" because I've heard before "accident", I've never heard about incident before, but because it was mentioned two or three times, I thought ok, that must be right. I just know the word accident but not the word incident.

O: Any other word combinations?

Interviewee Z: No... It was firefighter but that's probably because I don't know this word. I know what it means but what it could mean. to me it sounds a little bit strange. Fireman I would say but not a firefighter.

O: What about the multiple-choice exercise?

Interviewee Z: Sometimes problem to translate the sentence. [Reads Sentence 4] Does that mean smoking helped his voice to improve? What does it mean?

Appendix G: Post-test Survey. Stage I

Part 1

1. Your age: _____

2. Your gender (please circle): Male   Female

3. Is English your first language? (please circle) Yes   No

If you answered "yes" in Question 3, please proceed to <u>Questions 4 - 7</u> and skip Questions 8- 13.

If you answered "no" in Question 3, please proceed to <u>Questions 8 - 13</u> and skip Questions 4 - 7.

4. Please name another language except English (if any) that you know best:

_____

5. Your knowledge of the language named in **Question 4** is:

    a)   As good as your knowledge of English

    b)   Worse than your knowledge of English

6. How would you rate your knowledge of the language named in **Question 4** on the scale from 1 (*poor*) to 5 (*excellent*)?

7. At what age did you begin learning the language named in **Question 4**?

_____

8. What is your first language?

9. Your knowledge of English is:

    a)   As good as your knowledge of the language named in **Question 8**

    b)   Worse than your knowledge of the language named in **Question 8**

10. How would you rate your knowledge of English on the scale from 1 (poor) to 5 (excellent)?

11. At what age did you begin learning English?

12. How many years/months have you been studying English?

13. How many years/months did you live and study English in an English-speaking country (for example, Canada, Great Britain, the USA)?

Part 2

1. How fast did you read? (circle one)

    a) Slower than usual

    b) At my normal speed

    c) Faster than usual

2. What did you focus your attention on while you were selecting incorrect word combinations? (circle one)

    a) Word meaning

    b) Word form (for example, parts of speech, grammar, and so on)

    c) Both

    d) Other (please explain): _____

3. While you were reading, you mostly focused on (circle one):

    a) Word-by-word individually

    b) Combinations of two-three words

    c) Combinations of five or more words

    d) Sentences

Part 3

1. What language do you mostly communicate in?

    a) English

    b) Other (please name): _____

    c) Both

If in Part 1 you indicated English as your second language, please answer <u>Questions 2 and 3</u>, and skip Question 4.

If in Part 1 you indicated English as your primary language, please answer <u>Question 4</u> and skip Questions 2 and 3.

2. To what extent does your first language influence your English? In a few words explain how.

       1   2   3   4   5

  (Never)   (Very rarely)   (Sometimes) (Often)   (Very often)

3. How do you prefer to learn new words in English:

    a) One-by-one, not necessarily related to one another

    b) In sentences/combinations with other words

    c) Other (please name)

---

4. To what extent does another language that you know/learn influence your English? In a few words explain how.

       1   2   3   4   5

  (Never)   (Very rarely)   (Sometimes) (Often)   (Very often)

Appendix H: Test Items. Stage II

	Adjective + Noun	Verb + Noun	Verb + Adjective	Adverb + Verb	Adverb + Adjective
High F - High MI	Ethnic minority (F = 10.1, MI = 10.07) Wide range (F = 42.5, MI = 9.88) Vast majority (F = 11.31, MI = 11.03)	Provide evidence (F = 8.44, MI = 5.94) Collect data (F = 5.52, MI = 8.62) Play a role (F = 14.23, MI = 9.04)	Make explicit (F = 11.71, MI = 6.67) Become aware (F = 6.15, MI = 7.51)	Fully understand (F = 9.38, MI = 7.18) Widely used (F = 10.23, MI = 6.14)	Mutually exclusive (F = 5.56, MI = 13.24) Slightly different (F = 9.6, MI = 7.35)
Low F - Low MI	Limited access (F = 1.53, MI = 5.12) Present evidence (F = 1.29, MI = 3.2) Clear structure (F = 1.12, MI = 3.03)	Start the process (F = 1.12; MI = 3.75) Give access (F = 1.03, MI = 3.76) Take an approach (F = 2.11, MI = 3.58)	Make available (F = 3.59, MI = 4.2) Consider relevant (F = 1.08, MI = 4.48)	0	Particularly appropriate (F = 1.57, MI = 4.59) Relatively common (F = 1.17, MI = 4.25)
Low F - High MI	Stark contract (F = 1.26, MI = 10.17) Internal conflict (F = 1.8, MI = 8.08) Disposable income (F = 1.62, MI = 12.66)	Draw attention (F = 2.51, MI = 9.16) Cast doubt (F = 1.75, MI = 10.07) Raise awareness (F = 1.26, MI = 9.32)	Remain constant (F = 1.57, MI = 6.56) Prove successful (F = 2.56, MI = 8.81)	Heavily rely (F = 1.93, MI = 10.29) Adversely affect (F = 1.88; MI = 11.56)	Highly variable (F = 2.69, MI = 7.3) Broadly similar (F = 2.29, MI = 7.99)

Appendix I: Confidentiality Agreement

Confidentiality Agreement between Pearson Education Limited and Olga Makinina

CONFIDENTIALITY AGREEMENT

BETWEEN

1) **Pearson Education Limited** trading as Language Testing, a company registered in England and Wales under company number **872828** whose registered office is at **80 Strand, London WC2R 0RL,** (**"Pearson Education"**); and

2) **Olga Makinina,** (**"Recipient"**).

THE PARTIES AGREE as follows:

1. **Definitions and Interpretation**

1.1 For the purposes of this agreement the definitions in this clause shall apply.

"Confidential Information"	has the meaning given to it in clause 2.1.
"Disclosing Party"	means the party to this agreement that has disclosed the Confidential Information which, in the case of Pearson Education, shall include any member of its Group.
"Group"	means a party and each of its holding undertakings and subsidiary undertakings and associated companies.
"Purpose"	means the project in which the "recipient" is involved: a doctoral study using the items of the Academic Collocation List (ACL) for the purpose of developing a collocation-recognition-focused test and teaching tasks
"Recipient"	means the party to this agreement that has obtained the Confidential Information.

1.2 References in this agreement to:

(a) any enactment, or statutory instrument will be construed to include any subsequent amendments or re-enactments of them or their replacements;

(b) a person includes a reference to a body corporate, association or partnership and includes that person's successors.

1.3 Terms defined in the Companies Act 2006 shall, unless the context otherwise requires, have the same meanings in this agreement.

2. **Confidential Information**

2.1 Confidential Information means:

(a) information in whatever form obtained or maintained (including written, oral, visual and electronic) relating to the Disclosing Party or in any way concerning the business undertakings

1

Mutual Disclosure

and operations of the Disclosing Party which is obtained from time to time for the Purpose by or on behalf of the Recipient, its associates or subsidiary undertakings and/or by any of their respective directors, employees, advisors or agents; and

(b) analyses, compilations, studies and other documents or materials which contain or otherwise reflect or are derived or generated from any information described in clause 2.1(a).

2.2 Notwithstanding clause 2.1, Confidential Information shall not include any information which:

(a) is or becomes publicly available other than as a result of the information being disclosed in breach of this agreement;

(b) the Recipient can prove was known to the Recipient prior to the date of this agreement and the Recipient is not under any obligation of confidence in respect of the information; or

(c) is or has been lawfully received from a third party without restriction as to its use or disclosure.

3. Obligations

3.1 In consideration of the Disclosing Party disclosing Confidential Information the Recipient undertakes to the Disclosing Party that it will:

(a) treat and safeguard as private and confidential all the Confidential Information;

(b) use the Confidential Information only for the Purpose; and

(c) not disclose the Confidential Information to anyone or make copies of it except as permitted by this agreement. The data will be used only for statistical or research purposes. The data will not be published on line, in journals, books or by any means without previous agreement with the Disclosing Party. In any case, only a subset of the items may be reproduced and made available in the public domain (after agreement with the Disclosing Party).

4. Disclosure

4.1 The Recipient may disclose Confidential Information only:

(a) to such of its employees or officers (or, in the case of Pearson Education, the employees or officers of another member of its Group) as are strictly necessary for the Purpose;

(b) to professional advisers engaged to advise the Recipient;

(c) to the extent required by the laws or regulations of any country with jurisdiction over the affairs of the Recipient (or in the case of Pearson Education, any member of its Group) or any order of any court of competent jurisdiction or any competent judicial, governmental or regulatory body; or

(d) if it has been approved for release or use (in either case without restriction) by written authorization of the Disclosing Party.

4.2 The Recipient shall procure that persons to whom Confidential Information is disclosed in accordance with this agreement are made aware of the provisions of this agreement and observe the same restrictions on the use of the Confidential Information as are contained in this agreement.

4.3 Where the Recipient is required to disclose Confidential Information in accordance with clause 4.1(c), the Recipient shall give to the Disclosing Party reasonable prior notice of the need to comply with such requirement and shall have regard to the views of the Disclosing Party as to the

Mutual Disclosure

manner of compliance.

5. Copies

5.1 The Recipient may make only such copies of the Confidential Information as are necessary for the Purpose and shall clearly mark all such copies as confidential. All Confidential Information shall be returned to the Disclosing Party or destroyed, promptly upon request, together with all copies made by the Recipient.

6. No Warranty or Obligation to Disclose

6.1 The Disclosing Party makes no representation, warranty or assurance (express or implied) as to the accuracy, correctness or completeness of any Confidential Information or in the case of Confidential Information that is the subject of any intellectual property rights, that it has any such rights to or in the Confidential Information.

6.2 Nothing in this agreement shall be interpreted to transfer any rights in or require the disclosure of any Confidential Information.

7. Term and Termination

7.1 This agreement shall be effective from the date of its execution by both parties and shall continue in full force and effect for five years from the date of last disclosure of Confidential Information or termination of this agreement.

8. Inadequacy of Damages

8.1 The Recipient acknowledges and agrees that damages would not be an adequate remedy for any breach by it of this agreement and that, without prejudice to all other remedies the Disclosing Party may be entitled to as a matter of law, the Disclosing Party shall be entitled to the remedies of injunction, specific performance and other equitable relief for any threatened or actual breach of the provisions of this agreement and no proof of special damages shall be necessary for the enforcement of this agreement.

9. Relationship

9.1 This agreement shall not be construed as a joint venture, agency, partnership or other arrangement and is solely for the purpose of protecting Confidential Information. Nothing in this agreement grants to either party any right to make commitments of any kind for or on behalf of the other party without the prior written consent of the other party.

10. No Assignment

10.1 This agreement may not be assigned or transferred by either party without the other party's express written consent.

10.2 The recipient acknowledges that the Confidential Information relates to Pearson Education Group as a whole and agrees that Pearson Education or any member of Pearson Education Group may enforce the undertakings in this agreement.

11. Waiver

11.1 No failure by a party (or any member of its Group) to exercise, nor any delay in exercising any right, power, privilege or remedy under this agreement shall constitute a waiver by that party of any such

3

Mutual Disclosure

346

right, power or privilege, nor shall any single or partial exercise thereof preclude any further exercise of any such right, power or privilege.

12. Governing Law and Jurisdiction

12.1 This agreement shall be governed by and construed in accordance with English law and each party submits to the jurisdiction of the courts of England in respect of disputes relating to or in connection with this agreement.

The parties agree to the above as signified by executing below:

Signed for and on behalf of
Pearson Education Limited

Signature:

Name:
Veronica Benigno

Date:
15th January 2016

Signed for and on behalf of
Recipient

Signature:

Name:
Olga Makinina

Date: 18th January 2016

4

Mutual Disclosure

Appendix J: Matrix of Independent Test Variables. Stage II

Dependent Variable	Independent Variables		
Test Score (n = 34)	**Variable 1. Frequency (F)**	**Variable 2. Mutual information (MI)**	**Variable 3. Morphosyntactic types (MSTs)**
	Low F - Different MI - Different MSTs (n = 22)	High MI - Different F - Different MSTs (n = 24)	*Variable 3.1:* Adjective + Noun - Different F - Different MI (n = 9)
			Variable 3.2: Verb + Noun - Different F - Different MI (n = 9)
	Combinations of Factors of F + MI		*Variable 3.3:* Verb + Adjective - Different F - Different MI (n = 6)
	1. High F - High MI - Different MSTs (n = 12)		
	2. Low F - Low MI - Different MSTs (n = 10)		*Variable 3.4:* Verb + Adverb - Different F - Different MI (n = 4)
	3. Low F - High MI - Different MSTs (n = 12)		*Variable 3.5:* Adverb + Adjective - Different F - Different MI (n = 6)

Combinations of Factors of F + MI + MST

1. *Adjective + Noun*
   A. High F - High MI (n = 3)
   B. Low F - Low MI (n = 3)
   C. Low F - High MI (n = 3)

Dependent Variable	Independent Variables
	D. Low F - Different MI (n = 6)
	E. High MI - Different F (n = 6)
	2. Verb + Noun
	A. High F - High MI (n = 3)
	B. Low F - Low MI (n = 3)
	C. Low F - High MI (n = 3)
	D. Low F - Different MI (n = 6)
	E. High MI - Different F (n = 6)
	3. Verb + Adjective
	A. High F - High MI (n = 2)
	B. Low F - Low MI (n = 2)
	C. Low F - High MI (n = 2)
	D. Low F - Different MI (n = 4)
	E. High MI - Different F (n = 4)
	4. Verb + Adverb
	A. High F - High MI (n = 2)
	B. Low F - High MI (n = 2)
	5. Adverb + Adjective
	A. High F - High MI (n = 2)
	B. Low F - Low MI (n = 2)
	C. Low F - High MI (n = 2)
	D. Low F - Different MI (n = 4)
	E. High MI – Different F (n = 4)

Appendix K: Collocation Recognition and Controlled-production Test. Stages II and III

INSTRUCTIONS

There are 34 questions in this vocabulary test. Each question is a sentence in which some words are missing.

Sample question:

1. Please read this sentence.

In order to start a business in Cambodia, a business owner needs to have licenses from different governmental organizations and _____.

2. Use the table below this sentence to fill in the blank. Choose and circle only <u>one</u> <u>word from each column: one word from column A, <u>one word from column B, and so</u> <u>on.</u>

A	B
narrow	dominance
local	power
neighborhood	authority

All three words in one column have **similar meaning** (that is, they are synonyms). You need to decide which of the three synonyms fits the sentence best.

The words *narrow, local,* and *neighborhood* have similar meaning in this sentence: "restricted to a particular area".

The words *dominance, power,* and *authority* also have similar meaning in this sentence: "the right to control and make decisions".

Answer:

In order to start a business in Cambodia, a business owner needs to have licenses from different governmental organizations and <u>local authority.</u>

3. Please do NOT use dictionaries. All the word meanings have been provided for you in the test.

<u>Thank you for your participation in this research!</u>

TEST

1. While most students were Caucasian, the largest _____ of students reported themselves as Hispanic.

A	B
ethnic	adolescence
native	minority
traditional	opposition

The words *ethnic, native, and traditional* have similar meaning in this sentence: "related to a particular cultural group or region".

The words *adolescence, minority, and opposition* have similar meaning in this sentence: "a group of people that is usually discriminated against or does not have the same rights/responsibilities as other people".

2. Teachers have a _____ of duties from working with individual students to preparing classroom materials.

A	B
loose	mixture
spacious	range
wide	rank

The words *loose, spacious, and wide* have similar meaning in this sentence: "expansive;

includes a large amount of something or covers a large area".

The words *mixture, range, and rank* have similar meaning in this sentence: "a set of

similar things; series".

3. The _____ of companies with the new

workplace health and wellness programs reported reduced sickness absence.

A	B
spacious	majority
wide	mass
vast	plurality

The words *spacious, wide, and vast* have similar meaning in this sentence: "expansive;

includes a large amount of something or covers a large area".

The words *majority, mass, and plurality* have similar meaning in this sentence: "the

greater part/number".

4. _____ to the Internet prevents students from being able to find

necessary information for their research paper.

A	B
confined	access
finite	door
limited	entrance

The words *confined, finite, and limited* have similar meaning in this sentence: "restricted

in size, amount, or extent".

The words *access, door, and entrance* have similar meaning in this sentence: "a means of getting near, at, or to something".

5. The _____ in the research about technology suggests that males use digital technology more than females.

A	B
present	sign
prompt	substantiation
topical	evidence

The words *present, prompt, and topical* have similar meaning in this sentence: "of immediate relevance, happening now, related to current events".

The words *sign, substantiation, and evidence* have similar meaning in this sentence: "something that shows/indicates something clearly".

6. The study of time management reports that it takes a lot of time to get students organized. Another problem is the lack of _____ of the lesson.

A	B
clear	order
lucid	skeleton
sharp	structure

The words *clear, lucid, and sharp* have similar meaning in this sentence: "understandable; apparent".

The words *order, skeleton, and structure* have similar meaning in this sentence: "the way something is built or arranged".

7. Nearly one in three (30%) men did not know they could check themselves for cancer. This is in _____ women, who were more likely to take their health into their own hands.

A	B	C
blunt	contrast	to
entire	diversity	into
stark	opposite	towards

The words *blunt, entire, and stark* have similar meaning in this sentence: "absolute; complete".

The words *contrast, diversity, and opposite* have similar meaning in this sentence: "difference".

8. In the movie "Catastrophe", citizens decide between helping their loved ones and running for safety. This _____ is reflected in their movements and face expression.

A	B
enclosed	clash
internal	combat
intestine	conflict

The words *enclosed, internal, and intestine* have similar meaning in this sentence: "situated on the inside".

The words *clash, combat, and conflict* have similar meaning in this sentence: "fight or disagreement".

9. Costs associated with living in different geographical regions affect the

family's _____.

A	B
disposable	gains
consumable	income
usable	pay

The words *disposable, consumable, and usable* have similar meaning in this sentence:

"expendable; available".

The words *gains, income, and pay* have similar meaning in this sentence: "earnings".

10. The findings of the study of peer evaluation _____ that

students have the capability to assess and reward each other's performance

accurately.

A	B
cater	deposition
equip	evidence
provide	manifestation

The words *cater, equip, and provide* have similar meaning in this sentence: "supply;

give".

The words *deposition, evidence, and manifestation* have similar meaning in this sentence:

"proof; available information confirming that something is true".

11. A test checklist was used to _____ on the number of questions

each participant answered correctly.

A	B

collect	data
congregate	dossier
stockpile	score

The words *collect, congregate, and stockpile* have similar meaning in this sentence: "accumulate; gather".

The words *data, dossier, and score* have similar meaning in this sentence: "facts, statistics, and other information collected together".

12. Women _____ in keeping cultural traditions and values.

A	B	C
operate	a	role
personate	the	portrayal
play	-	appearance

The words *operate, personate, and play* have similar meaning in this sentence: "occupy oneself in activity".

The words *role, portrayal, and appearance* have similar meaning in this sentence: "a part that someone has in a particular situation or activity".

13. Children adopted from abroad usually arrive home with developmental delays in their first language and then have

to _____ learning a second language.

A	B	C	D
arouse	a	measure	about
depart	the	process	from
start	-	progress	of

The words *arouse, depart, and start* have similar meaning in this sentence: "begin doing something from a particular point in time or space".

The words *measure, process, and progress* have similar meaning in this sentence: "method; series of actions to achieve result".

14. Universities are prohibited to _____ personal information about students.

A	B	C
contribute	access	to
give	approach	into
supply	path	for

The words *contribute, give, and supply* have similar meaning in this sentence: "offer; provide".

The words *access, approach, and path* have similar meaning in this sentence: "a means of getting near, at, or to something".

15. The purpose of this article is to _____ the ways that the use of digital technology is integrated into the learning standards.

A	B	C
convey	attention	to
convince	concentration	into
draw	contemplation	for

The words *convey, convince, and draw* have similar meaning in this sentence: "move; evoke; influence".

The words *attention, concentration, and contemplation* have similar meaning in this sentence: "the act of close or careful observing".

16. Since her previous way of engaging students was unsuccessful, Lisa planned to _____ that students never have time to misbehave.

A	B	C
grasp	an	access
seize	the	approach
take	-	entrance

The words *grasp, seize, and take* have similar meaning in this sentence: "get; help oneself to".

The words *access, approach, and entrance* have similar meaning in this sentence: "a means of getting near, at, or to something".

17. Studies conducted with different groups _____ whether the children and adolescents benefit from the authoritative style of parenting.

A	B	C
aim	confusion	over
cast	distrust	at
direct	doubt	on

The words *aim, cast, and direct* have similar meaning in this sentence: "send something in a particular direction".

The words *confusion, distrust, and doubt* have similar meaning in this sentence: 'lack of confidence; questioning".

18. School counselors can engage faculty and staff in activities that reduce prejudice and _____ their own and students' cultural selves.

A	B	C
aggravate	acquaintance	on
exalt	awareness	about
raise	realization	of

The words *aggravate, exalt, and raise* have similar meaning in this sentence: "increase".

The words *acquaintance, awareness, and realization* have similar meaning in this sentence: "knowledge of someone or something".

19. The specialists ensure that the information they _____ for the public is comprehensible and accessible.

A	B
do	available
make	obtainable
produce	procurable

The words *do, make, and produce* have similar meaning in this sentence: "cause something to be".

The words *available, obtainable, and procurable* have similar meaning in this sentence: "easy to get".

20. Most museums do not want to _____ where their collections come from.

A	B
do	outspoken

make	unequivocal
produce	explicit

The words *do, make, and produce* have similar meaning in this sentence: "cause something to be".

The words *outspoken, unequivocal, and explicit* have similar meaning in this sentence: "specific; stated clearly and in detail; leaving no doubt".

21. Students need to be educated about gender stereotypes so that they can _____ their own biases related to gender roles.

A	B	C
become	alert	on
convert	awake	from
mature	aware	of

The words *become, convert, and mature* have similar meaning in this sentence: "to grow/evolve into something".

The words *alert, awake, and aware* have similar meaning in this sentence: "knowledgeable".

22. Government agencies use all knowledge and information that they _____ to the particular situation.

A	B
consider	allowable
ponder	pointful
scrutinize	relevant

The words *consider, ponder, and scrutinize* have similar meaning in this sentence: "regard/view something or someone in a certain way; think about something".

The words *allowable, pointful, and relevant* have similar meaning in this sentence: "suitable or right".

23. There are many differences between traditional and nontraditional learning environments, but the role and importance of class discussion _____.

A	B
endure	constant
remain	interminable
persist	unfluctuating

The words *endure, remain, and persist* have similar meaning in this sentence: "stay in the same place or in the same condition".

The words *constant, interminable, and unfluctuating* have similar meaning in this sentence: "not changing or varying; continuous".

24. The clinical studies reported that cancer treatment strategies _____ when tested on laboratory animals.

A	B
confirm	blooming
justify	notable
prove	successful

The words *confirm, justify, and prove* have similar meaning in this sentence: "show that something is true; establish facts".

The words *blooming, notable, and successful* have similar meaning in this sentence: "achieving the results wanted or hoped for".

25. It seems that once people_____ what a panic attack is, they are no longer frightened of the physical symptoms.

A	B
fully	deduce
intimately	surmise
utterly	understand

The words *fully, intimately, and utterly* have similar meaning in this sentence: "completely".

The words *deduce, surmise, and understand* have similar meaning in this sentence: "perceive/know the intended meaning".

26. IELTS scores are _____ to select students for college admission.

A	B
thickly	operated
vastly	used
widely	expended

The words *thickly, vastly, and widely* have similar meaning in this sentence: "to a great extent".

The words *operated, used, and expended* have similar meaning in this sentence: "deployed; implemented".

27. Universities and libraries _____ the Internet to share resources to students across the country through online content.

A	B	C
count	densely	in
entrust	heavily	at
rely	thickly	on

The words *count, entrust, and rely* have similar meaning in this sentence: "depend on; have confidence in".

The words *densely, heavily, and thickly* have similar meaning in this sentence: "to a great degree; with a lot of force or effort".

28. The disturbance in the classroom might

_____ children's concentration and time on task.

A	B
adversely	alter
resentfully	affect
unsympathetically	touch

The words *adversely, resentfully, and unsympathetically* have similar meaning in this sentence: "in a way that is harmful and likely to cause problems".

The words *alter, affect, and touch* have similar meaning in this sentence: "influence; cause change".

29. The statistics showed that race and ethnicity were

not _____ categories. On the contrary,

race and ethnicity complement and influence one another.

A	B
cooperatively	exclusive
jointly	sole
mutually	unique

The words *cooperatively, jointly, and mutually* have similar meaning in this sentence:

"done by two or more people in the same way; together".

The words *exclusive, sole, and unique* have similar meaning in this sentence: "limited;

restricted to one person only or a group of people".

30. When your essay looks wrong, you edit the notes, making them more coherent.

The next day, as you come at the writing with a fresh eye, the notes look wrong

again, but this time for _____ reasons.

A	B
faintly	different
inconsiderably	disparate
slightly	various

The words *faintly, inconsiderably, and slightly* have similar meaning in this sentence: "to

a small degree; a little".

The words *different, disparate, and various* have similar meaning in this sentence:

"unlike; not the same".

31. Teachers try to select _____ materials to interest their students.

A	B
decidedly	appropriate
markedly	congruous
particularly	opportune

The words *decidedly, markedly, and particularly* have similar meaning in this sentence: "especially; more than usual".

The words *appropriate, congruous, and opportune* have similar meaning in this sentence: "suitable or right".

32. The research suggests that some minor aggressive behaviors are _____ in children, but serious rule violations are abnormal during childhood.

A	B
approximately	common
proportionately	daily
relatively	habitual

The words *approximately, proportionately, or relatively* have similar meaning in this sentence: "by comparison".

The words *common, daily, and habitual* have similar meaning in this sentence: "ordinary; normal; often happening".

33. The company indicated that production costs are _____ and are influenced by the employees' experience and training.

A	B
highly	mutable
supremely	shifting
vastly	variable

The words *highly, supremely, and vastly* have similar meaning in this sentence: "to a large degree; very".

The words *mutable, shifting, and variable* have similar meaning in this sentence: "changeable".

34. The pattern of mortality is _____ for men and women.

A	B
broadly	collateral
grandly	same
widely	similar

The words *broadly, grandly, and widely* have similar meaning in this sentence: "generally; largely".

The words *collateral, same, and similar* have similar meaning in this sentence: "related; not different".

Appendix L: Post-test Survey. Stages II and III

Please answer the following questions:

1. What is your age? (circle one)

a. 18 - 20

b. 21 - 25

c. 26 - 30

d. More than 30

2. What is your mother tongue (first language)?

3. Do you speak any other languages except your mother tongue and English? If yes,

please list these languages.

4. At what age did you start learning English?

5. How many years/months have you been learning/speaking English?

6. How many years/months have you been living in an English-speaking country?

7. How do you prefer to learn new words in English? (circle one)

    d) One-by-one, not necessarily related to one another

    e) In sentences/combinations with other words

    f) Other (please indicate) :

8. Please rate this test on the scale from 1 (very easy) to 5 (very difficult).

1	2	3	4	5
Very easy	Somewhat easy	Neither too easy nor too difficult	Somewhat difficult	Very difficult

Appendix M: Collocations Used for Teaching Module at Stage III

Ethnic minority

Wide range

Vast majority

Limited access

Present evidence

Clear structure

Stark contrast

Internal conflict

Disposable income

Provide evidence

Collect data

Play a role

Start the process of

Give access to

Draw attention to

Take the approach

Cast doubt on

Raise awareness of

Make available

Make explicit

Become aware of

Consider relevant

Remain constant

Prove successful

Fully understand

Widely used

Rely heavily on

Adversely affect

Mutually exclusive

Slightly different

Particularly appropriate

Relatively common

Highly variable

Broadly similar

Appendix N: Teaching Module

<u>*Notes*</u>

The rationale and description of the teaching activities can be found in Chapter 7.

This is an abridged version of exercises.

Lesson 1

Objectives: Students will be introduced to the concept of collocations; will learn to recognize collocations; will identify collocation functions; and will be able to identify links between collocational components

Time: 40 minutes

Materials: the BNC and COCA corpora

Sequence of tasks:

Activities	Type of work	Rationale	Time frame
Exercise 1. Comparing correct and incorrect collocations	Groups of three-four	Introduce collocations; discuss the importance of using correct (that is, natural-sounding/native-like) word combinations	5 minutes
Exercise 2. Finding collocates for frequent polysemantic verbs	Groups of three-four	Consolidate the concept of collocations and learn about differences between collocations, free word combinations, and other	15 minutes

		formulaic units	
Exercise 3. Noticing and finding synonyms	Individual/pair-share	Explain collocation functions: clarity and fluency of communication	10 minutes
Exercise 4. Combining words into collocations and matching parts of collocations	Individual homework	Explain the links between collocation components	10 minutes

Excerpts from exercises:

Exercise 1

Decide which of these word combinations are correct. Why do you think so?

See clearly   Play a role

See transparently   Operate a role

Do available   Internal conflict

Make available   Intestine conflict

Working definition: *Collocations are fixed combinations of two or more words. If we change one of the words for another word, which has similar meaning to the original word, it will not sound English-like.*

Exercise 2

Please think of as many words as you can that could combine with the following verbs:

Take   Do

Make   Keep

Give   Break

Collect

Exercise 3.

Please read these sentences and complete the following tasks:

A. Find and underline two-word collocations; and

B. Change these collocations into synonymous (similar in meaning) words or word combinations.

(1). People living in rural areas have limited access to health care services.

(2). In order to end an international or internal conflict, negotiations often must be held with the leaders who are responsible for war crimes and crimes against humanity.

(3). The woman's dark-colored dress is a stark contrast to her vividly red shoes.

(4). Although the thief swore he had not stolen anything, the large sum of money in his pocket cast doubt on his sincerity.

(5). Most people do not fully understand that advertisement plays an important role in their lives.

(6). Cybersecurity and privacy are not mutually exclusive.

(7). The weather was highly variable, with temperatures rising to 30 degrees in full sun and then dropping 20 degrees under cloud cover.

Exercise 4

1. *Please combine these words into two (or three)-word collocations:*

Affect   Become   Draw   Provide   Rely   Use   Attention

Evidence   Aware   Adversely   Heavily   Widely   Of   On

To

2. Match two parts of sentences so as to have correct collocations:

The question is not clear. Please repeat it in a **slightly**	**similar** requirement.
Since I have problems with time management, my friend advised me to attend the time management workshop before the finals. That was a **particularly**	**affect** performance.
The two jobs have a **broadly**	**attention** to them.
The high level of stress could **adversely**	**different** way.
My vision is poor, so I cannot **clearly**	**appropriate** advice.
The cooperation among the project team members can be made easier if the leading manager **makes** the project objectives	**constant**.
Admit your weaknesses but do not d**raw**	**see** letters.
Terrorism and crime **remain**	**explicit.**

Lesson 2

Objectives: Students will explore the morphosyntactic collocation structure (a node and collocates); will learn to differentiate and use five basic morphosyntactic types of collocations (Adjective + Noun, Adverb + Adjective, Verb + Adjective, Verb + Adverb,

and Verb + Noun); will learn how to use online collocation dictionaries and databases;

and will be able to guess the node/collocate(s) making inferences from the context

Time: 40 minutes

Materials: the BNC and COCA corpora; Pro-Writing Aid: Free Online Collocation

Dictionary; JustTheWord (JTW). Online Collocation Database

Sequence of tasks:

Activities	Type of work	Rationale	Time frame
Exercise 1. Underlining nodes: (a) verbs, (b) nouns, (c) adjectives; and collocates: (a) nouns, (b) adjectives, (c) adverbs in sentences	Pairs/groups of three-four	Identify morphosyntactic collocational types; learn the rules of combining words	15 minutes
Exercise 2. Using online collocation dictionaries and databases to find collocates for a set of nodes	Individual work; pair/share	Consolidate understanding of the five morphosyntactic types; identify links between collocational components	10 minutes
Exercise 3. Filling	Groups of four +	Increase awareness	15 minutes

374

in the blanks in concordance lines	homework	of the five morphosyntactic types; find nodes based on contextual clues	

Excerpts from exercises:

Exercise 1

See Lesson 1, Exercise 3, for the list of sentences with collocations. Read underlined two (or three)-word combinations and answer the following questions:

- What parts of speech do these collocations consist of?
- What is the main word in each collocation?

Exercise 2

1. *Read these words and identify what parts of speech they belong to. Decide what parts of speech these words can combine with:*

Majority   Similar

Income   Exclusive

Provide

Start

Become

Prove

Use

Rely

2. *Go to this website: http://www.just-the-word.com/*

Type the following main words in the search box; click on COMBINATIONS; write out

four supporting words for each main word.

Majority [search for ADJECTIVES ONLY]:

Income [search for ADJECTIVES ONLY]:

Provide [search for NOUNS ONLY]

Start [search for NOUNS ONLY]

Become [search for ADJECTIVES ONLY]

Prove [search for ADJECTIVES ONLY]

Use [search for ADVERBS ONLY]

Rely [search for ADVERBS ONLY]

Similar [search for ADJECTIVES ONLY]

Exclusive [search for ADJECTIVES ONLY]

Exercise 3

Read these sentences and try to guess a missing word for each group. What part of

speech will this word belong to?

A.

A number of services are available for women from ethnic _____ communities.

Telephone advice services are available in various ethnic _____ languages.

Social changes in Britain have resulted in a varied ethnic _____population.

B.

The student is able to present ideas in a clear _____ both verbally and written.

The new security system offers a crystal-clear _____ and includes data

security, thanks to digital data storage.

This essay has a logical clear _____.

C.

Researchers _____data for a range of surveys.

The second phase of the study will need to _____data.

To know how many people live on a minimum wage, we need to _____data for

the past five years.

D.

A week of screenings, debate and discussion to _____awareness of the United

Nations' work.

Students decided to work on a project to _____

awareness of the causes of poverty and injustice in the Third World.

This video _____awareness of India's orphans.

E.

All employees and must _____ aware of the changes that have to be made to their

working habits.

 Attending WTS/LSS workshops helps participants _____aware of their writing and

reading problems.

Political leaders must _____ aware of people's problems.

F.

This proposal is useless because it means abandoning a good policy that _____

successful.

There is no vaccine for the prevention of the disease and no treatment _____ successful so far.

This bonus _____ successful in encouraging all employees to come to work on time.

G.

The driver did not fully _____ how operate his new car.

When people get married, they need to fully _____ commitments to each other.

Customers fully _____ what they want to buy.

H.

We _____ heavily on the generosity of our sponsors.

Teenagers may _____ heavily on support from their peers when their parents get divorce.

Many countries _____ on the citizen's taxes.

I.

It seems to me that it is a particularly _____ time for us to take a few moments to think about our future.

Online marketing is particularly _____ for people who cannot go to the stores.

Working abroad is particularly _____ for young people.

J.

The terms and conditions in the new contract are broadly _____ to the original contract.

Both politicians came to an agreement because they reached broadly

_____ conclusions about the types of changes they want to make in the

society.

Broadly _____tuition fees will be paid for the fall and winter terms.

Lesson 3

Objectives: Students will learn to recognize and use collocations with the low and high

frequency of occurrence and low and high mutual information score

Time: 45 minutes

Materials: the BNC and COCA corpora; Just-The-Word (JTW), Pro-Writing Aid: Free

Online Collocation Dictionary

Activities	Type of work	Rationale	Time frame
Exercise 1. Using online collocation dictionaries to determine collocates that most and least frequently combine with a set of nodes	Individual/pair-share	Introduce low - high frequency of collocations and their components (nodes and collocates)	10 minutes
Exercise 2. Finding a node for a group of collocates and determining their	Groups of three-four	Check the participants' ability to differentiate between low and	10 minutes

		high frequency collocations	
Exercise 3. Dividing collocations into weaker and stronger word combinations	Individual/pair-share	Introduce low and high strength of association between collocational components	10 minutes
Exercise 4. Building associative networks	Groups of three-four	Increase the participants' awareness of the frequency and strength of association links between collocational components	5 minutes
Exercise 5. Filling in the blank	Individual homework	Develop the participants' ability to use contextual clues and increase their awareness of semantic links between the words	10 minutes

Excerpts from exercises:

Exercise 1

See Lesson 2, Exercise 2, for the list of nodes. Look at the four collocates that you have found for each node. Using online collocation databases (for example, http://www.just-the-word.com/), find frequencies for each collocation. Which collocations are more or less frequent?

Exercise 2

Read these words and try to predict what the main word for each group of supporting words could be. Using online collocation databases (for example, http://www.just-the-word.com/), find frequencies for each supporting word when combined with the main word. Which collocations are more or less frequent?

A.	B.	C.
Broad	Attract	Widely
Wide	Draw	Regularly
Full	Occupy	Frequently
Whole	Capture	Primarily

Exercise 3

1. *Read these collocations and divide them into two sets: STRONGER and WEAKER.*

Disposable income

Clear structure

Cast doubt on

Raise awareness of

Take the approach

Start the process of

Become aware of

Consider relevant

Rely heavily on

Widely used

Mutually exclusive

Highly variable

Relatively common

2. *Look at the scores on the scale of strength of relationship between words that was assigned to each collocation and discuss the differences/similarities between this score and your sets of STRONGER and WEAKER collocations:*

Collocation	Score on the scale of strength of relationship between words
Disposable income	12.66
Clear structure	3.03
Cast doubt on	10.07
Raise awareness of	9.32
Take the approach	3.58
Start the process of	3.75
Become aware of	7.51
Consider relevant	4.48
Rely heavily on	10.29

Widely used	6.14
Mutually exclusive	13.24
Highly variable	7.3
Relatively common	4.25

Exercise 4

Think about and write down as many words as you can that could combine with these main words. Using online collocation databases (for example, http://www.just-the-word.com/), find additional supporting words and their frequencies. Count the number of supporting words for each main word. Collocations with which main words would be the strongest?

Range (Noun)

Minority (Noun)

Cast (Verb)

Affect (Verb)

Exclusive (Adjective)

Similar (Adjective)

Exercise 5

Fill in the blank in each sentence by following these steps:

(a) Select the main word from this list:

Appropriate; Become; Exclusive; Give; Make; Provide; Structure; Range; Understand

(b) Try to guess the supporting word based on the main word and the meaning of the whole sentence.

(1). When writing an academic paper, you need to _____ to support your main idea.

(2) The university _____ to the library to all students.

(3) The paper got a high mark because it had a very _____.

(4) Students at the university have a _____ of tasks and assignments.

(5) The grades will be _____ to the students at the end of the term.

(6) When writing papers, students should _____ of the writing conventions in English.

(7). Could you please repeat the question? I am afraid I cannot _____ what you mean.

(8) Human rights and bullying are _____.

(9) Feeling overwhelmed is _____ for the first-year undergraduate students.

Lesson 4

Objectives: Students will be able to recognize and correct incorrect collocations and use collocations when speaking and writing

Time: 40 minutes

Materials: the BNC and COCA corpora

Activities	Type of work	Rationale	Time frame
Exercise 1. Recognizing and correcting incorrect	Pair/ groups of three-four	Develop "inner feeling" based on understanding how	10 minutes

collocations		links between collocational components work	
Exercise 2. Writing a story and filling in the blanks using the list of 34 collocations	Individual homework; pair/share	Memorize and review 34 academic collocations	15 minutes
Exercise 3. 4-3-2-minutes-long talk using the list of 34 collocations	Groups of 4	Memorize and review 34 academic collocations	15 minutes

Excerpts from exercises:

Exercise 1

Please find, underline, and correct two consecutive word combinations that seem to you wrong or odd-sounding:

1. Native minority groups might often live in rural areas or conflict-affected areas.

2. Many Canadians have finite access to the essential health resource of Pharmacare.

3. The Canadian government branches have a sharp structure.

4. The intestine conflict in the country will result in violation of human rights and freedoms.

5. An employer may require an employee who takes an extended medical leave to equip evidence such as doctor's note.

6. The doctors will not supply access to any information about their patients.

7. The evidence from the witnesses cast confusion on the driver being responsible for the car accident.

8. The teacher made outspoken that students should prepare for the final exam.

9. Any information or evidence may be considered allowable to the crime scene.

10. The threat of terrorism and crime remains unfluctuating in the coming years.

11. English is vastly used around the globe.

12. The high level of stress will resentfully affect your health.

13. The final in-class writing is faintly different from the midterm writing.

14. Flooding is a proportionately common occurrence in Thailand.

15. Access to safe water and sanitation is vastly variable across the region, especially in rural areas and for the poorest people.

Appendix O: Post-instruction Test. Stage III

Notes:

The number of questions and test format remain the same as in the pre-instruction test (see Appendix K).

The sentences in the test and their order of items in the collocational matrix were changed to eliminate the factor of preliminary memorization of responses.

INSTRUCTIONS

There are 34 questions in this vocabulary test. Each question is a sentence in which some words are missing.

Sample question

1. Please read this sentence.

Example: In order to start a business in Cambodia, a business owner needs to have licenses from different governmental organizations and **_____**.

2. Use the table below this sentence to fill in the blank. Choose and circle only <u>one</u> word from each column: one word from column A, <u>one word from column B, and so on.</u>

A	B
narrow	dominance
local	power
neighborhood	authority

All three words in one column have **similar meaning** (that is, they are synonyms). You need to decide which of the three synonyms fits the sentence best.

The words *narrow, local,* and *neighborhood* have similar meaning in this sentence: "restricted to a particular area".

The words *dominance, power,* and *authority* also have similar meaning in this sentence: "the right to control and make decisions".

Answer:

In order to start a business in Cambodia, a business owner needs to have licenses from different governmental organizations and <u>local authority.</u>

3. Please do NOT use dictionaries. All the word meanings have been provided for you in the test.

TEST

1. 42% of students attending public schools are considered to be an

_____.

A	B
native	minority
ethnic	opposition
traditional	adolescence

The words *native, ethnic, and traditional* have similar meaning in this sentence: "related to a particular cultural group or region".

The words *minority, opposition, and adolescence* have similar meaning in this sentence: "a group of people that is usually discriminated against or does not have the same rights/responsibilities as other people".

2. Good stories demonstrate a _____ of characters and problems.

A	B
wide	rank
spacious	range
loose	mixture

The words *wide, spacious, and loose* have similar meaning in this sentence: 'expansive; includes a large amount of something or covers a large area".

The words *range*, *rank, and mixture* have similar meaning in this sentence: "a set of similar things; series".

3. The _____ **of speech in class consisted of teacher talk.**

A	B
wide	plurality
vast	mass
spacious	majority

The words *wide, vast, and spacious* have similar meaning in this sentence: "expansive; includes a large amount of something or covers a large area".

The words *plurality, mass, and majority* have similar meaning in this sentence: "the greater part/number".

4. Many children in developing countries have _____ **to education and could be forced into child labor.**

A	B
limited	door
finite	entrance

confined	access

The words *limited, finite, and confined* have similar meaning in this sentence: "restricted in size, amount, or extent".

The words *door, entrance, and access* have similar meaning in this sentence: "a means of getting near, at, or to something".

5. _____ indicates that people can inherit left-handedness from their parents.

A	B
topical	evidence
present	sign
prompt	substantiation

The words *topical*, *present, and prompt* have similar meaning in this sentence: "of immediate relevance, happening now, related to current events".

The words *evidence, sign, and substantiation,* have similar meaning in this sentence: "something that shows/indicates something clearly".

6. The _____ of the textbook - reading passages and exercises - means that students will find it easy to read.

A	B
clear	skeleton
lucid	structure
sharp	order

The words *clear, lucid, and sharp* have similar meaning in this sentence: "understandable; apparent".

The words *skeleton, structure, and order* have similar meaning in this sentence: "the way something is built or arranged".

7. Most people prefer to live in wealthy cities that exist in

_____ the poverty of rural areas.

A	B	C
stark	opposite	into
entire	diversity	to
blunt	contrast	towards

The words *stark, entire, and blunt* have similar meaning in this sentence: "absolute; complete".

The words *opposite, diversity, and contrast* have similar meaning in this sentence: "difference".

8. To avoid _____ , people's behavior should be consistent with their beliefs and values.

A	B
enclosed	conflict
internal	combat
intestine	clash

The words *enclosed, internal, and intestine* have similar meaning in this sentence: "situated on the inside".

The words *conflict, combat, and clash* have similar meaning in this sentence: "fight or disagreement".

9. People who live in middle- and upper-class neighborhoods can ask neighbors and friends to donate their _____ to the lower-class neighborhoods.

A	B
consumable	income
disposable	gains
usable	pay

The words *consumable, disposable, and usable* have similar meaning in this sentence: "expendable; available".

The words *income, gains, and pay* have similar meaning in this sentence: "earnings".

10. The rhythm and depth of our breathing _____ of our emotions.

A	B
cater	evidence
equip	manifestation
provide	deposition

The words *cater, equip, and provide* have similar meaning in this sentence: "supply; give".

The words *evidence, manifestation, and deposition* have similar meaning in this sentence: "proof; available information confirming that something is true".

11. Teachers _____ on student progress every week.

A	B
stockpile	data

congregate	dossier
collect	score

The words *stockpile, congregate, and collect* have similar meaning in this sentence: "accumulate; gather".

The words *data, dossier, and score* have similar meaning in this sentence: "facts, statistics, and other information collected together".

12. Teacher attitudes about their students' ability may

_____ in the students' capacity to learn.

A	B	C
operate	a	appearance
personate	the	role
play	-	portrayal

The words *operate, personate, and play* have similar meaning in this sentence: "occupy oneself in activity".

The words *appearance, role, and portrayal* have similar meaning in this sentence: "a part that someone has in a particular situation or activity".

13. Many countries tried to _____ adapting to climate change.

A	B	C	D
arouse	a	process	from
depart	the	measure	of
start	-	progress	about

393

The words *arouse, depart, and start* have similar meaning in this sentence: "begin doing something from a particular point in time or space".

The words *process, measure, and progress* have similar meaning in this sentence: "method; series of actions to achieve result".

14. The university health centre is required to

_____ all basic medical treatments.

A	B	C
contribute	path	for
give	access	into
supply	approach	to

The words *contribute, give, and supply* have similar meaning in this sentence: "offer; provide".

The words *path, access, and approach* have similar meaning in this sentence: "a means of getting near, at, or to something".

15. The World Heritage List is designed to

_____ the wealth and diversity of Earth's cultural and natural heritage.

A	B	C
convey	contemplation	for
draw	concentration	into
convince	attention	to

The words *convey, draw, and convince* have similar meaning in this sentence: "move; evoke; influence".

The words *contemplation, concentration, and attention* have similar meaning in this sentence: "the act of close or careful observing".

16. The law safeguards a person's interests regardless of whether he is rich or poor, weak or strong. The United States needs to _____ to its international leadership.

A	B	C
grasp	an	entrance
take	the	approach
seize	-	access

The words *grasp, take, and seize* have similar meaning in this sentence: "get; help oneself to".

The words *entrance, approach, and access* have similar meaning in this sentence: "a means of getting near, at, or to something".

17. Many patients feel as though they cannot ask questions or

_____ their doctors' opinions.

A	B	C
aim	confusion	at
cast	doubt	on
direct	distrust	for

The words *aim, cast, and direct* have similar meaning in this sentence: "send something in a particular direction".

The words *confusion, doubt, and distrust* have similar meaning in this sentence: "lack of confidence; questioning".

18. The health centers at schools _____ the

dangers of smoking.

A	B	C
exalt	realization	at
aggravate	awareness	of
raise	acquaintance	on

The words *exalt, aggravate, and raise* have similar meaning in this sentence: "increase".

The words *realization, awareness, and acquaintance* have similar meaning in this

sentence: "knowledge of someone or something".

19. The library considers which e-books they would like to

_____ to the students.

A	B
do	procurable
make	obtainable
produce	available

The words *do, make, and produce* have similar meaning in this sentence: "cause

something to be".

The words *procurable, obtainable, and available* have similar meaning in this sentence:

"easy to get".

20. Students were able to _____ what they knew, what

they could do, how they learned, and what was important about their learning.

A	B
do	explicit

make	outspoken
produce	unequivocal

The words *do, make, and produce* have similar meaning in this sentence: "cause

something to be".

The words *explicit, outspoken, and unequivocal* have similar meaning in this sentence:

"specific; stated clearly and in detail; leaving no doubt".

21. Students need to _____ the traditions, values, and

customs of various cultures.

A	B	C
convert	aware	from
become	awake	on
mature	alert	of

The words *convert, become, and mature* have similar meaning in this sentence: "to

grow/evolve into something".

The words *aware, awake, and alert* have similar meaning in this sentence:

"knowledgeable".

22. Students need to learn how to ask questions about whatever they

_____ for their coursework.

A	B
ponder	relevant
consider	pointful
scrutinize	allowable

The words *ponder, consider, and scrutinize* have similar meaning in this sentence: "regard/view something or someone in a certain way; think about something".

The words *relevant, pointful, and allowable* have similar meaning in this sentence: "suitable or right".

23. The school policy might change; however, teachers and students will

_____ .

A	B
remain	unfluctuating
endure	interminable
persist	constant

The words *remain, endure, and persist* have similar meaning in this sentence: "stay in the same place or in the same condition".

The words *unfluctuating, interminable, and constant* have similar meaning in this sentence: "not changing or varying; continuous".

24. The same advertising methods that

_____ **in influencing one customer could just as easily be ignored by another.**

A	B
confirm	successful
justify	notable
prove	blooming

The words *confirm, justify, and prove* have similar meaning in this sentence: "show that something is true; establish facts".

The words *successful, notable, and blooming* have similar meaning in this sentence: "achieving the results wanted or hoped for".

25. Students _____ **course requirements and are prepared to succeed.**

A	B
fully	surmise
intimately	understand
utterly	deduce

The words *fully, intimately, and utterly* have similar meaning in this sentence: "completely".

The words *surmise, understand, and deduce* have similar meaning in this sentence: "perceive/know the intended meaning".

26. Taking notes while listening to a lecture is an important strategy that students _____ **for increasing attention and retaining content.**

A	B
expend	thickly
operate	widely
use	vastly

The words *expend, operate, and use* have similar meaning in this sentence: "deployed; implemented".

The words *thickly, widely,* and *vastly* have similar meaning in this sentence: "to a great extent".

27. Colleges and universities _____ student feedback in terms of classroom performance of teachers.

A	B	C
count	heavily	at
rely	thickly	in
entrust	densely	on

The words *count, rely, and entrust* have similar meaning in this sentence: "depend on; have confidence in".

The words *heavily, thickly, and densely* have similar meaning in this sentence: "to a great degree; with a lot of force or effort".

28. Exposure to air pollution may _____ the newborn babies.

A	B
resentfully	affect
unsympathetically	alter
adversely	touch

The words *resentfully, unsympathetically, adversely* have similar meaning in this sentence: "in a way that is harmful and likely to cause problems".

The words *affect, alter, and touch* have similar meaning in this sentence: "influence; cause change".

29. Family and career might not be _____ . One can learn to balance both.

A	B
cooperatively	sole
mutually	unique
jointly	exclusive

The words *cooperatively, mutually, and jointly* have similar meaning in this sentence:

"done by two or more people in the same way; together".

The words *sole, unique, and exclusive* have similar meaning in this sentence: "limited;

restricted to one person only or a group of people".

30. Students have written three research papers; each paper provides

_____ information on grammar and

punctuation problems.

A	B
inconsiderably	different
faintly	various
slightly	disparate

The words *inconsiderably, faintly, and slightly* have similar meaning in this sentence: "to

a small degree; a little".

The words *different, various, and disparate* have similar meaning in this sentence:

'unlike; not the same".

31. Debates on "hot" topics are _____ for

teachers who strive for students' total involvement.

A	B
particularly	opportune

markedly	congruous
decidedly	appropriate

The words *particularly, markedly, and decidedly* have similar meaning in this sentence: "especially; more than usual".

The words *opportune, congruous, and appropriate* have similar meaning in this sentence: "suitable or right".

32. Grade 10 represents an age when alcohol use is

_____ among youths.

A	B
relatively	habitual
proportionately	common
approximately	daily

The words *relatively, proportionately, and approximately* have similar meaning in this sentence: "by comparison".

The words *habitual, common, and daily* have similar meaning in this sentence: "ordinary; normal; often happening".

33. Symptoms of Lyme disease are _____ and

may include fever, headache, fatigue, and skin rash.

A	B
supremely	variable
highly	shifting
vastly	mutable

The words *supremely, highly, and vastly* have similar meaning in this sentence: "to a large degree; very".

The words *variable, shifting, and mutable* have similar meaning in this sentence: "changeable".

34. The UK and the US have _____ historical backgrounds in relation to their awareness of child abuse.

A	B
grandly	similar
widely	collateral
broadly	same

The words *grandly, widely, and broadly* have similar meaning in this sentence: "generally; largely".

The words *similar, collateral, and same* have similar meaning in this sentence: "related; not different"