The Correlation Between Teaching Collocations and Lexical Bundles and the Improvement in the Writing Skill of First-Year University Students

Submitted by Sally Kondos
to the University of Exeter as a thesis for the degree of
Doctor of Education
May 2023

I certify that all material in this thesis which is not my own work has been identified and that any material that has previously been submitted and approved for award of a degree by this or any other University has been acknowledged.

ABSTRACT

This study explores the correlation between the teaching of collocations and lexical bundles and the improvement of the writing skill of first-year university students. The thesis addresses three research questions. First, to what extent can the explicit teaching of collocations and lexical bundles assist the students in learning them and can later use the acquired collocations and lexical bundles morphologically and grammatically accurately in writing? Second, the study asks if there is a correlation between using the collocation and lexical bundle and improving the writing grade. Finally, the study investigated if there is a relationship between the increase in the number of collocations and lexical bundles and the difference in the writing grade and whether such an increase leads to improvement in the overall score. It is worth mentioning that the collocations and lexical bundles are two of the leading representatives of the Formulaic language. Many studies discussed the significant role that collocations and the lexical bundles play in helping English as a foreign language students (EFLs) to express themselves accurately in writing, besides providing them with the knowledge to produce coherent and precise text. The study implemented quantitative research, and the findings were the outcome of the statistical analysis of the pre-test and post-tests and written assignments of the control and experimental groups. The findings concluded that the explicit instruction of the collocation and lexical bundles significantly improved the overall writing grade of the experimental group. The study identified some commonly used collocations and lexical bundles among the high-achieving participants, and the number of the collocations and lexical bundles used in writing positively correlated with an improved overall grade. More suggestions will be discussed on how more attention should be given to incorporating the collocation and lexical bundles in the English for Academic Purposes programmes in universities.

ACKNOWLEDGEMENTS

I want to dedicate my doctorate to the beautiful soul of my father, George Kondos. If you are looking down upon me, which I am sure you are, I want to tell you that I did keep my promise of finishing my doctorate. I love you, Dad, and I hope I made you proud.

I owe a big thank you to my wonderful kids, Tommy and Taline. You have accompanied me with enthusiasm, love, and care during this journey. Thank you for constantly encouraging me to finish my studies. To Amir, my husband, none of this could be possible without your support.

Finally, I extend my gratitude to the University of Exeter for providing me with this exciting opportunity. I also want to thank my supervisors, Dr. Esmaeel Abdollazadeh and Dr. Philp Durrant, for their dedication and support throughout the process. Thanks to Dr. Philip Durrant for always guiding me and for all his commitment and publications that helped me to find my passion.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. Introduction

The chapter introduces the study's context regarding the setting, the participants, and the background of the study. The chapter discusses the significance of the research questions and the importance of the topic in bridging the gap in the teaching of formulaic language, particularly collocations and lexical bundles, and improving students' writing skills in English academic programs.

1.1 Background of the Study

The study took place in the context of Dubai, in the United Arab Emirates (UAE). According to the rules of the Higher Ministry of Education in the UAE, all non-native English students who aspire to learn in any English-speaking Western university must pass a standardized test to start their undergraduate studies. If they do not acquire the required grade, whether on the Test of English as. Foreign Language (TOEFL), the International English Language Testing System (IELTS), or the university Accuplacer exam, they are enrolled in an English Bridge Program (ENGB) for a year to enhance their language proficiency level in all skills. In my teaching context, most students struggled to score a 6.5 out of 9 in the writing part of the IELTS exam or a 550 out of 667 in the TOEFL test. Students' main weakness lies in their inability to elaborate their ideas and develop their writing further; they struggle to expand on a sub-topic in the writing exam. As a way of helping the students in the bridge programme, the teachers constantly advise students to memorize more words to have vocabulary knowledge instead of working on different strategies for building better sentences structure that can assist them in developing their ideas further (Kondos, 2020).

To help ENGB students, teachers at the university where the study took place provide weekly vocabulary lists adopted from the Academic Word List (AWL). However, over my 9 years working in ENGB, I have not seen any benefits from using these single-word vocabulary lists. This is simply because students memorize each word without studying its different aspects, like what part of speech the word is or its collocations, which eventually ends in failing to use the words accurately in writing. Consequently, the students memorize words without being able to use them in their own sentences.

The current study proposed to focus more on formulaic language with a focus on collocations and lexical bundles rather than a single word. As early as the 1980s, several linguists reasoned against the Chomskyan approach, which suggested that any natural language is the outcome of a set of countless utterances that are mostly generated from a group of grammatical structures (Pawley & Syder, 1983). According to Barlow (2000), Liu and Huo (2011), and Wray (2000), language users are inclined to use particular words in their language production more than going through the hassle of forming complicated grammatical structures that could express the same concept. The expanding interest in what is known as the formulaicity of language production has been linked to the growing emphasis on new linguistic theories that focus on performance rather than on competence. This interest emphasizes the substantial role of formulaic sequences in language production (Al Hassan & Wood, 2015).

1.2 The Rationale of Implementing Collocations and Lexical Bundles

There are many language categories under the umbrella of formulaic language, such as spoken idioms (Liu, 2003), phrasal verbs (Garnier & Schmitt, 2015), academic collocations (Achermann & Chen, 2013), and lexical bundles. Many studies have investigated how English-language teachers can incorporate formulaic language in their classes (Schmitt, 2022).

The current study investigated the impact of collocations and lexical bundles as representative of formulaic language since the research aims to examine the correlation between the explicit teaching of collocations and lexical bundles and the improvement in the writing skills of first-year university students. The rationale for using collocations & lexical bundles originated from the idea that mastering language is characterized by the repetition of formulaic patterns such as fixed and semi-fixed multi-words, which is the case of collocations and lexical bundles (Byrd & Coxhead, 2010). Applied linguistics scholars have argued that the ability to recognize patterns, recall them, and later reproduce them contributes greatly to language proficiency and to fluency in mastering the English language. The focus on emphasizing the direct relationship between the collocations and lexical bundles and the improvement in the L2 writing skills stemmed from the rationale that

language production, especially writing, is not the result of a word-word formation that is governed by the syntactic rule. Alternatively, writing is a product of formulaic languages such as collocations and lexical bundles because such sequences of words are retrieved from the memory as chunks rather than separate individual words. (e.g., Baker, 2006; Biber & Barbieri, 2007; Biber et al., 2004; Cortes, 2004; Hyland, 2008). Kormos (2006) argued that learning collocations and lexical bundles can accelerate written language processing because these forms contribute to fluency processing due to the use of the prefabricated chunks of language expressions that add to the clarity development of the L2 written texts (Chen, 2012).

It is essential to clearly distinguish between the collocation and lexical bundles because each formulaic language will be introduced and taught differently, as mentioned in Table 2. Collocations and lexical bundles are crucial representatives of the formulaic language. However, each has its characteristics; hence taught differently. First, the collocations occur within a certain distance and with high frequency; they occur in different language combinations; for example, verb + noun collocations – *get attention to*, *leave a message*, verb adverb collocation – *badly damaged*, *strongly advise*, adverb + verb collocation – *fully understand, fully appreciated*; and finally verb+ adjective collocation – *quite good, highly profitable*(Granger, 2018).

On the other hand, lexical bundles are high-frequency combinations of words that usually occur next to each other. They are incomplete grammatical structures formed by combining two noun phrases, verb, and noun phrase, or more phrases and clauses. The lexical bundles are building blocks of language. They do not have any syntactic integrity and cannot stand on their own; they must be incorporated and integrated within the sentences. They are classified into three different types; referential bundles – *at the end of – at the beginning of – in the interpretation of,* discourse organizing bundles – *on the other hand*, and finally, attitudinal lexical bundles – *it should be noted that* (Dontcheva-Navratilova,2012).

1.3 Significance of the Study

The significance of the study lies in its importance to the English Bridge Program (ENGB) students. The students are all Arabic-language natives, making

them English as foreign language learners (EFLs). It is worth mentioning that our ENGB students typically encounter teaching methods that focus on routine memorization and translation of English teaching materials into the students' mother tongue (Celce-Murcia, 1991).

During my nine years teaching in the ENGB program, and after following some of the student's progress, I noticed that the ENGB students spend three months and sometimes a year in the program with no significant improvement in their writing skills, and this affects their overall university experience, as the studies are all in English.

For this reason, I developed the study to promote the importance of writing in the ENGB. The collocations refer to lexical items that occur within a particular linear frequency, such as *write an essay*. The lexical bundles refer to the high-frequency combinations of words that can appear next to each other, such as, *on the other hand,* and *you are responsible* (Durrant, 2017). Using collocations and lexical bundles can intrigue students to learn more about the English language, and the knowledge of the formulaic language through collocations and lexical bundles can open the door to a new perspective of language teaching. It can train students to be proactive in their writing instead of simply memorizing lists of vocabulary words without being able to use them accurately in sentences. Using collocation and lexical bundles in the ENGB can help students improve their sentence structure and create more fluent and coherent writing. By developing their writing skills, they will have a better learning experience during their university years, especially since the teaching medium is all English throughout their 4-year university education.

The significance of the study goes beyond the betterment of the English language teaching practices in the ENGB bride program. The significance of the current study is raising awareness and encouraging more and more educators to investigate the impact of collocations and lexical bundles in their teaching context. Only through longitudinal research and published papers are more educators in foreign language studies informed about the impact of lexical bundles and collocations on improving foreign language fluency, whether spoken or written. It is worth mentioning that the successful implementation of collocations and lexical bundles is recommended in any foreign language pedagogy, not necessarily the

English language. The lexical bundles and collocations play a significant role in Arabic as a foreign language programmes in universities worldwide. On the use of collocation and lexical bundles in Arabic as foreign language curricula, Arabe(2010) argued that Formulaic language, with its two representatives, collocations, and lexical bundles, has become a common practice in taught language research. The rationale behind the increased emphasis on teaching collocations and lexical bundles comes from the many benefits of learning collocation and lexical bundles. Among many, the collocational patterns of learning any language will train the learners to look at any language as a string of words rather than a single stand-alone word that they often fail to incorporate into meaningful sentences. Another benefit, the knowledge of lexical bundles and collocation saves much time; due to automated language production, the learners have more time to focus on the content they plan to develop and express(Arabe, 2010)

1.4 Research Questions

The research questions that guided this study were the following:

1. To what extent do the explicit teaching of collocations and lexical bundles help students use them accurately?
2. Is there a correlation between the use of collocation and lexical bundles and the improvement of students' overall writing grades?
3. Is there a relationship between the increase in the number of lexical bundles used and the change in essay scores after intervention?

CHAPTER 2. Literature Review

The chapter introduces the empirical research on the impact of teaching collocations and lexical bundles to improve the writing skill of EFLs. The chapter discusses the existing research and publications that served as a background to the current research. The literature review gives a complete account of how collocations and lexical bundles emerged from vocabulary studies to become two of the most researched features of teaching formulaic language.

The main themes in this literature review include the importance of vocabulary in language acquisition, followed by the required vocabulary size of the EFL learners to be considered competent in English, and the difference between receptive and productive vocabulary knowledge.

The literature review will discuss how to measure the depth and breadth of vocabulary knowledge and how words do not exist in isolation, followed by the different methods of measuring the vocabulary size of English as Foreign language learners. After discussing the significant role of vocabulary knowledge in learning English, the literature review will discuss the impact of the formulaic language on second language acquisition. Since the study's main aim is writing skills, it is imperative to review how formulaic teaching language can improve writing skills and the processing of formulaic language.

It is also essential for the study to highlight how the formulaic language bridges the gap in English for Academic Purposes programmes. A whole section will discuss the relationship between formulaic language and students' progress in academic writing; the section will also introduce the shift from single-word to multiword knowledge. The literature review will discuss the rationale for choosing collocations and lexical bundles as two representatives of the formulaic language representations. Finally, the challenges of developing tests to examine the knowledge of collocations and lexical bundles.

2.1 The Importance of Vocabulary in Language Acquisition

Vocabulary is part of the core of mastering a foreign language. Improving language fluency has always been associated with increasing vocabulary knowledge (Adolphs & Schmitt, 2004). While grammar and vocabulary are

considered the two pillars of acquiring the English language, scholars and educators have emphasized the importance of vocabulary over grammar, arguing that, without vocabulary, students cannot express themselves much (Thornbury, 2002). It is undeniable that vocabulary is a potent tool in language acquisition. A language learner is considered fluent or competent in English if they are capable of expressing themself with abundant expressions. Richard and Renandya (2002) argued that vocabulary plays a crucial role in learners' language proficiency; if learners fail to build a solid language knowledge, they find it difficult to communicate successfully in English. Richard and Renandya (2002) noted that, "[v]ocabulary is a core component of language proficiency and provides much of the basis for how well learners speak, listen, read, and write. Without an extensive vocabulary and strategies for acquiring new vocabulary, learners often achieve less than their potential and may be discouraged from making use of language learning opportunities around them such as listening to radio, listening native speakers, using the language in different context, reading, or watching TV" (p. 258).

Despite the growing interest in vocabulary today, teaching vocabulary was somehow ignored in the past. During the 1960s, the focus was on the audio-lingual language teaching approach; researchers at that time believed that the emphasis should be on teaching grammar and phonology because, once these aspects were mastered, vocabulary could be easily acquired (Kurniawan, 2016). The demand for teaching vocabulary increased in 1980 due to the growing use of computers in language learning, marking the beginning of the communicative approach. The advocates of the communicative approach argued that mastering vocabulary comes at an early stage of language acquisition, and then grammar follows.

The emphasis on vocabulary teaching increased as linguists argued that an adequate number of vocabulary words is essential for successful communication (Kurniawan, 2016). Recently, English as a foreign language program developers have realized the importance of vocabulary in language proficiency and no longer consider vocabulary as supplementary materials to the curriculum. Instead, syllabus designers now regard vocabulary as an essential tool that needs further development for language development (Kurniawan, 2016). Many aspects contribute to increasing vocabulary knowledge. Studies have always emphasized

vocabulary's crucial role in mastering the English language's communicative skills. It is essential to understand that vocabulary does not mean memorizing words by heart; on the contrary, it is the ability to understand the various functions of words.

Lexical fluency and the ability to express oneself through meaningful lexis are the cornerstones of mastering a language. The ability to express oneself with ideas, feelings, and emotions by using accurate lexis distinguishes the competent user from the foreigner to any language. Learning a new word entails mental engagement that is far more complicated than simply memorizing words. Most teaching practices focus on increasing vocabulary rather than promoting better practices to know how to use the new word. This is a common criticism of applied linguistics advocates, who argue that word knowledge is unusually multi-faceted and requires the learners to be aware of the meaning and usage of words (Nation, 2001).

Language research long focused on the syntactical and morphological aspects of learning vocabulary. The assumption was based on the notion that vocabulary development is linear(Meara,2002). Little attention was given to lexical development in second-language acquisition. However, this changed in the early 2000s with research into lexical development and its impact on language proficiency (Meara, 2002). Researchers also began to argue that language learners should be introduced to a large amount of vocabulary at the beginner's level. They continue to argue that if a learner does not have at least 2,000 high frequency words as a start, they will continue to struggle to communicate (Murcia, 2001). This signals the importance of vocabulary size in acquiring the English language.

Before discussing the breadth and depth of vocabulary, it is essential to understand that vocabulary knowledge varies between receptive and productive vocabulary. It is worth noting that not all learned vocabulary can be considered productive knowledge, which makes productive words fewer in numbers than receptive words.

2.1.1 Vocabulary Size

Research on language studies have investigated the amount of vocabulary that EFL learners should acquire to be considered competent users of the English

language (Miralpiex & Muñoz, 2018). Research has also investigated the correlation between vocabulary knowledge, exemplified by vocabulary size, and proficiency in a particular language (Miralpiex & Muñoz, 2018). Meara (1996) argued that, “all other things being equal, learners with large vocabularies are more proficient in a wide range of language skills than learners with smaller vocabularies, and there is some evidence to support the view that vocabulary skills make a significant contribution to almost all aspects of L2 proficiency” **(**p.7).

Although this has been an ongoing topic, there is no definite answer to the question of vocabulary size. Up to a point, language scholars proposed various numbers for vocabulary size. Some suggested it would be helpful to start by identifying an adequate goal of vocabulary size for language learners, suggesting that 1114000-word families (Goulden et al., 1990) and 88500-word families (Nagy & Anderson, 1984) might be enough. However, it is challenging to plan language goals to help the learners gain such a large vocabulary. It is worth mentioning that vocabulary size is based on word families (Hu & Nation, 2000; Laufer, 1989; Nation, 2006).

Various attempts were made to find an adequate vocabulary size for EFL learners. Nation (2006) argued that the most common 1,000-word families have six numbers on average. Conversely, the number of word families decreases to three at the 9,000-frequency level (Nation, 2006). Based on Nation’s (2006) calculations, Schmitt (2008) suggested that if a learner has 6,000-word families for effective listening, the number would be equivalent to 28,015 individual word knowledge, where the knowledge of 8,000-word families for successful reading is equivalent to the knowledge of 34,660 different word forms.

As far as reading skill is concerned, Laufer **(**2004) argued that vocabulary size is essential to comprehending a text in English, with an estimation of 95% coverage of the text for the learner to understand the text, whereas Hu and Nation (2000) argued that a good percentage for successful comprehension without assistance was 98%. Hu and Nation’s (2000) assumption resulted from a study on the relationship between text coverage and reading comprehension for non-native speakers. They examined the number of running words non-native learners could identify in a reading text. They prepared some texts, replaced some low-frequency words with unfamiliar words, and tested if the learners could understand the text

using multiple-choice reading questions and a written recap for reading comprehension. After further calculations, the study concluded that 98% text coverage is required to gain comprehension of any reading text which concludes that the reading comprehension of the text increases as the word knowledge increase.

The result of this study goes in line with the findings of a similar study by Carver (1994), who argued that if a text is easy, close to 0% of the words will be unknown to the students, whereas when the text is hard, 2% of the words or more will be unknown. In other words, when the difficulty level of the material is equivalent to the ability level of the learner, then around 1% of the words will be unknown (Carvers, 1994, cited in Nation, 2006). In a similar study, Adolphs and Schmitt (2003, 2004) investigated word families' coverage in the Cambridge and Nottingham Corpus of Discourse in English (CANCODE) and BNC spoken corpora. The CANCODE comprises five million words. Adolphs and Schmitt (2003) adopted a different approach to determining the coverage of the words, by merely counting the words that actually occurred in the corpus. They aimed to investigate the percentage of coverage that word families in the corpus provided and the adequate vocabulary size a learner needed to be able to communicate effectively in the English language. Surprisingly, the result was that sufficient vocabulary size determines the success in everyday spoken activities (Adolphs & Schmitt, 2003).

According to these studies, an EFL should know between 6,000-7,000-word families in spoken discourse and between 8,000–9,000 in written discourse to be considered a competent user of the English language (Mutlu & Kaşlıoğlu, 2016). Other scholars have argued that 10,000 words in English are essential to pass university entry exams (Hazenberg & Hulstun, 1996). However, scholars advise that language researchers should deal with such numbers with caution because EFLs' vocabulary changes as they forget some words, so their vocabulary size changes frequently (Meara & Rodriguez Sánchez, 1993).

Understandably, all of these figures are challenging to achieve in any language program. Schmitt (2010) explained that EFLs fail to acquire these numbers of word families because most language programs emphasize teaching individual words rather than their collocations and associations because they assume that single words are easy to teach and implement into their teaching

materials. Conversely, Lewis (1997) and Hill (2000) argued that if EFLs are presented with new individual words alongside co-occurring words, their vocabulary size would increase.

Research on the adequate vocabulary size of language has proven the role of lexical knowledge in acquiring a target language. Many studies have also investigated the relationship between competence in vocabulary and L2 writing. Although studies indicate the importance of lexical knowledge in L2 writing, there is not enough research to confirm the correlation between specific vocabulary size and the quality of writing (Miralpiex & Muñoz, 2018).

Studies by Read (2005) and Graham et al. (2009) attempted to investigate the relationship between vocabulary size and quality of writing. Graham et al. (2009) investigated the effect of a six-month intervention course on improving French students' English-productive written and spoken skills. The study's results confirmed the correlation between receptive vocabulary size and writing and listening scores. However, it also confirmed that there was no massive significance in the writing and listening test results. Read (2005) examined the speaking extracts of IELTS candidates in an attempt to analyze their lexical choices and diversity, and found that the richness of lexical choices led to higher IELTS grades.

2.1.2 Receptive and Productive Vocabulary Knowledge

The changing vocabulary size of EFL learners raises the question of what it means to know a word and the different aspects associated with identifying a word. In other words, it is not the vocabulary size that matters; it is the learner's ability to use the words successfully. Nation (2010) proposed that word knowledge means the learner can recognize the word in spoken and written form and can use it successfully in any written activity. However, other scholars have argued that Nation's (2010) definition implies that form and meaning are two separate entities of vocabulary knowledge. Besides, this definition ignores the other aspects of knowing a word beyond the form and meaning (Laufer et al., 2004), such as the receptive (passive) and active (productive) knowledge that scholars have recently used (Milton, 2009). However, the focus should be the learning of the word rather than receptive or productive knowledge (Shahov, 2012).

Receptive knowledge is associated with listening or reading, while productive knowledge is linked with speaking or writing (Laufer & Goldstein, 2004). The distinction is not simple, as a reader or listener requires passive skills to understand a text and guess the unfamiliar words, while also using productive skills to deal with the texts (Milton, 2009).

For Nation (2001), word knowledge is complicated because it includes knowledge of form, knowledge of meaning, and meaning of use. Nation (2001) also divided every category into a subcategory. For example, knowledge contains spoken, written, and word parts. The aim of the subcategorization is an attempt to answer how the learners learn a word. Understanding the form/spoken/receptive component helps learners understand how a word looks like while comprehending the form/spoken/productive aspects of the word helps the learner use the word successfully (Nation, 2001).

Before Nation (2001), Henriksen (1999) proposed that lexical knowledge has three components: partial to precise knowledge, shallow to deep knowledge, and receptive to productive knowledge. Henriksen's (1999) description was an attempt to unify lexical knowledge (Shahov, 2012). Scholars have continued to investigate the relationship between receptive and productive word knowledge. Meara (1997) argued that the two types of knowledge are entirely different entities and that there is no continuum between them, but Henriksen (1999) proposed that some lexical origination in which the productively known words are associated to a productive item, though receptively known words are not associated to any words in the lexicon. Despite the continuous attempts of language researchers to understand the dimensions of word knowledge, there is no clearly accepted definition of receptive and productive word knowledge (Shahov, 2012).

2.1.3 Measuring the Depth and Breadth of Vocabulary Knowledge

Building a strong vocabulary knowledge starts with the understanding that words do not exist in isolation, so every word must be studied in relation to other words. The focus should be on improving the lexical competence of learners rather than studying single words in isolation. Haastrup and Henriksen (2000) argued that lexical competence has three attributes: partial-precise, which is the ability to understand the different levels of the word; receptive-productive, and finally the

depth-of-knowledge, which means the ability to understand the relationship between a particular word and other words in the lexicon. In other words, this is the pragmatic and syntagmatic relationship between lexicon (Bardakçı, 2016). These scholars have shown that learning new words is far more complicated than merely memorizing them; it is a mental process that requires learners to make a mental connection between the lexicon. The ability to categorize lexical knowledge could present a unified lexical knowledge that is accepted worldwide. Language researchers thus no longer see vocabulary knowledge as a single entity, but rather as a multi-dimensional process with breadth and depth (Hatami & Tavakoli, 2013).

Measuring the depth of word knowledge, or how well a student has acquired a word, is far more complicated than measuring the breadth, which corresponds to the number of acquired words. The depth of word knowledge requires mastering several aspects of knowing a word, including semantic relationships, collocations, and syntactic patterning (Cobb, 1999). It is essential to build vocabulary knowledge depth because knowing a word is far more complicated than merely recalling its meaning (Ünaldı, 2011).

Admittedly, most common language programs adopt the developmental approach when considering word knowledge because they believe in the incremental nature of vocabulary and that learning occurs as a continuum. Adopting this approach makes designing tests challenging because no one can confirm how the word is learned, the different stages of learning a word, or how the learner has moved from one stage to another. Additionally, researchers interested in measuring the depth of vocabulary must note that this cannot happen in one single test because of the different elements of any word, such as spelling, register, or collocation (Schmitt, 2010).

The breadth of language refers to the number of words a learner can identify and define but not necessarily use in a sentence or in a communicative way. In other words, this type of vocabulary knowledge is linear and does not involve depth (Bardakçı, 2016).

Before testing vocabulary breadth, researchers must decide what lexical item to use as the measuring unit. As Schmitt (2010) stated, “Different ways of counting lexical items will lead to vastly different results, and a persistent problem

in lexical studies is that size figures are reported, but without a clear indication of how they were derived" (p. 188). (Shahov, 2012).

It is helpful to understand the differences between the lexical units that are commonly used in breadth tests. First, *tokens* are the number of running words that correspond to the number of words in a spoken or written text. Questions that assess tokens would ask about the number of words needed to read a book (Nation, 2010). Conversely, a *lemma* is the headword of any word and, according to psycholinguistics studies, the mind stores the headword (*teach*, *go*) but struggles to remember irregular forms (*taught*, *went*). Nation (2010) argued that when the lemma is considered a measuring unit in breadth tests, the number of words reduce significantly. For instance, the 61,805 types in the Brown corpus reduced to almost 37,617 lemma, which is about a 40% decrease (Nation, 2010).

Finally, *word families* cover all parts of speech or words related to the headword, such as *teach*, *taught*, *teacher*, *teachable* and *teaches*. Despite the similarity between word families and lemma, researchers argue that the lemma is the best lexical unit they facilitate the task of counting the productive types of vocabulary (Shahov, 2012).

Regardless of the opposing opinions on testing the depth and breadth of vocabulary knowledge, there have been promising conclusions that could lead to further studies on the depth of vocabulary. Because of the focus on the fluency and automaticity of language, more psycholinguistic techniques have been used to gain a better understanding of vocabulary knowledge (Schmitt, 2010, cited in Shahov, 2012). The continuous attempts to investigate different strategies to enhance EFLs' fluency in the English language have led to more research on speech fluency, which has focused only on the psycholinguistic foundation, with much attention to the automaticity through which L2 speech is retrieved and produced. Since formulaic sequences are considered chunks of language that are retrieved as a whole to enhance fluency, researchers have examined the impact of formulaic sequences on language fluency (Chen, 2012).

Laufer and Goldstein (2004) proposed automaticity or fluency as an additional aspect of vocabulary knowledge. Schmitt (2010) went to explain that fluency is the ability to understand what is being read or heard and to retrieve the

acquired knowledge to produce language successfully, whether written or spoken (Shahov, 2012).

2.1.4 The Measurement of Vocabulary Size

All research on vocabulary size needs to be validated by methods of measuring vocabulary. Such measuring tools help educators and English-language teachers determine the best teaching practices to promote vocabulary proficiency. Better vocabulary proficiency means a more competent user of the English language. Meera (1996) proposed that “learners with large vocabularies are more proficient in a wide range of language skills than learners with smaller vocabularies” (p. 37). English-language teachers always advise EFL learners to increase their vocabulary knowledge to be fluent in the language and use vocabulary tests to monitor the progress of their students. Although vocabulary tests might not lead to better teaching practices, they help the language teachers develop a better understanding of the learning process of EFL learners and come up with better ways to improve their teaching of the language (Shahov, 2012).

When discussing different vocabulary tests, the type of test adopted depends on the approach a researcher wants to pursue to assess the learners’ language progress. There are two types of procedures: the developmental and the dimensional approach. The tests developed under those approaches ask the test-taker to produce language in written or spoken text (Nation, 2010). The developmental approach tests the depth of vocabulary knowledge and regards vocabulary learning as a continuum, giving the incremental nature of vocabulary learning. It is challenging to use the developmental approach to design vocabulary tests. Despite all of the research in language learning, there are still questions regarding the number of stages learners take to acquire a new word, whether the stages are all equal in the lexical development, and the amount of time needed to learn a new word.

One of the vocabulary tests used to measure the vocabulary size of EFL learners is the Vocabulary Knowledge Scale (VKS), which is considered among the popular tests. However, one drawback of this test is the lack of unidimensional representation of lexical knowledge (Shahov, 2012). Schmidt (2010) suggested reducing the evaluation scale from five to four and added that the focus should be

on the learner's language proficiency and what they can do with the word rather than what the learner knows about the word. Schmidt (2010) added that there is no single vocabulary scale that can give a definite account of a lexical item's incremental route. It might be beneficial to look upon vocabulary testing from the perspective of a dimensional approach because lexical knowledge has multiple layers, such as spelling, register, and collocation. Researchers interested in testing vocabulary should consider the multiple layers of word knowledge and recognize that no single test can assess all the word knowledge layers. However, Schmitt (2010) advocates that such productive and passive tests are essential in understanding the nature of word knowledge and the relationship between every component. Importantly, it helps the examiner understand the path of the word from receptive to productive knowledge and whether this follows a continuum (Schmitt, 2010, cited in Shahov, 2012).

Other language researchers advise that there is no need to focus on testing the depth of vocabulary knowledge, as it "may be modified if it is to remain useful as a dimension at all since nothing keeps the elements comprising it together terribly persuasively and it does not seem to function entirely separately from breadth" (Milton, 2009, p. 169).

2.2 The Impact of Formulaic Language on Second-Language Acquisition

Formulaic sequences form a fundamental part of the English vocabulary (e.g., Martinez & Schmitt, 2012). They are looked upon as an essential element of how students learn the English language (e.g., Alali & Schmitt, 2012; Wood, 2002). It is far from possible to have a single exact definition of formulaic sequences. Wray (2002) proposed that more than 50 valid definitions exist. The term *formulaic sequence* covers a whole diversity of language expressions, from idioms and proverbs to collocations and ready-made chunks of expressions (Wray, 2002). These sequences function as a unit of language expression, even though they are composed of multiple words. Wray (2002) posited that, once they are mastered, they can be recalled from memory easily without being subject to any grammatical rules or structures.

On the other hand, researchers continue to argue that learning any lexeme takes time, comes gradually, and leads to various challenges, and that there is no

reason to think that learning formulaic sequences is any easier (Nation, 1990; Schmitt, 2000). Despite linguists' continuous reminders of the importance of integrating more formulaic sequence materials into the teaching materials of English as a second language, the debate on the effectiveness of teaching these sequences continues. In the following sub-sections, I discuss some studies that have examined whether formulaic sequences are acquired holistically and whether they continue to be a helpful tool in speeding up learning a second language or whether they are acquired incrementally like any other lexeme in English vocabulary.

2.2.1 Teaching Formulaic Language to Improve Writing Skills

The advantages of teaching formulaic sequences have been the subject of many recent studies (e.g., Durrant, 2008; Wray, 2000). For instance, Martinez and Schmitt (2012) noted some key advantages of formulaic sequences: they compose a large percentage of English discourse, whether written or spoken (e.g., 58.6% of speaking discourse and 52.3% of writing discourse; Erman & Warren, 2000), and they can facilitate the communication of many expressions (e.g., Conklin & Schmitt, 2008; Jiang & Nekrasova, 2007). Finally, they enable learners to produce language successfully by providing chunks of expressions (e.g., Guz, 2014). For all the above reasons, researchers are encouraged to pursue their investigation of better ways to integrate more formulaic sequences in classroom materials to promote more effective techniques of teaching English as a foreign language.

The significant results of many empirical studies in L1 discourse have encouraged educators to consider formulaic sequences as an auspicious tool that might promote better and more effective methods of teaching English as a foreign language (Boers & Lindstromberg, 2009; Jones & Haywood, 2004; Schmitt & Underwood, 2004). Linguists in the field of language acquisition advocate that teaching formulaic sequences to L2 learners might enable them to grasp the new language quickly and lead to more successful language production (Fitzpatrick, 2005; Osborne, 2008; Wood, 2000; Wray, 2002).

Researchers have also proposed that formulaic sequences should be considered a valuable tool in teaching academic writing, especially for L2 learners who experience difficulty in expressing themselves in writing (Hyland, 2003, 2006;

Leki, 2006; Reppen, 2002; Silva, 1993). The significance of formulaic sequences in validating English written skills can be attributed to the fact that they enable teachers to provide L2 learners with lists of formulas that they can use in their writing, and researchers argue that this practice helps learners develop into proficient writers (Coxhead & Byrd, 2007; Martinez & Schmitt, 2012). Lewis (1997) explained that, through frequent exposure to these lists and some more practice, L2 learners tend to use these expressions in their academic prose.

2.2.2 The Processing of Formulaic Language

Schmitt **(**1992) was among the applied linguists who drew attention to the importance of speech fluency, outlining fluency as an “automatic procedural skill.” Schmitt (1992) went on to explain that being fluent is the ability to produce speech effortlessly, whereas non-fluent speech requires a lot of attention and planning. Consequently, fluency depends on the “procedural knowledge” of a skilled learner instead of the “declarative knowledge” of a learner who consciously arranges vocabulary to produce speech (Schmitt, 1992, p. 6). This statement asks the fundamental question of whether language production is creative or memory-based.

Pawley and Syder (1983) argued that most produced language is not the outcome of a word-to-word formation governed by syntactic rules. On the contrary, it is produced through formulaic language, which are sequences of words and phrases retrieved from memory as a whole (Chen, 2012). Kormos (2006) advocated that formulaic language plays a vital role in the development of speech fluency because of the automatization of the encoding process and the use of prefabricated language units (Chen, 2012). All of this has placed formulaic sequences as one approach that can lead to accelerative language processing because formulaic sequences are part of a holistic system rather than an analytical one (Weinert, 1995).

Formulaic sequences are an essential part of all sorts of discourse, whether spoken or written (Kathy & Schmitt, 2008). They are not merely groups of words linked by collocation; on the contrary, they are often associated with a single meaning or pragmatic function, which explains their vital role in communication due to their pragmatic nature. Linguists have proposed that formulaic sequences

facilitate the processing of the target language because multi-word sequences are understood more quickly than non-formulaic words (Conklin & Schmitt, 2008). This preposition results from studies indicating that one-third to one-half of the English language is made of formulaic sequences (Foster, 2001).

2.2.3 Formulaic Language to Bridge the Gap in English for Academic Purposes (EAP)

Studies need to investigate the significance of the use of formulaic language in English for academic purposes (EAP). EAP refers to the teaching of English to university students. Language researchers advocate that formulaic language is essential to language acquisition because it supports fluency processing (Durrant, 2018). Formulaic language makes the job of learners easier because it provides them with plenty of ready-made chunks of phrases and language that the student can seek to understand, learn, and eventually produce (e.g., Cortes, 2006; Simpson-Vlach & Ellis, 2010). Ellis (2008) advocated that learning different language formulas plays a vital role in language acquisition because learners tend to learn grammatical structures by mastering formulas (e.g., Lewis, 2000; Nattigner & Decarrico, 1992). Linguists have also investigated the significance of learning formulas in accelerating the acquisition of target languages in general and for university students in particular. This interest is due to the idea that learning formulaic sequences might help university students be proficient in English.

Durrant (2018) urged that the ongoing attempts to link formulaic language to better performance in the language go back to Pawley and Syder's **(**1983) claim that the use of formulaic language makes the students perform "native-like" language production. Durrant (2018) explained that using the term "native-like" is very misleading because it assumes that because some learners were born in a particular community, this makes them privileged to speak in a certain way. Durrant (2018) noted that "native-like" can apply to any community, including the EAP community, because every community develops specific ways of communicating, which then take the form of formulas. The whole idea of learning the formulas is to build membership with those who are part of the community and who speak in a certain way (Wray, 2002).

The idea of membership is evident in EAP because the goal is to help learners become part of English-speaking communities. Fitting in and meeting the expectations of specific communities have been significant reasons for the associations of formulaic language with EAP. Studies have anticipated that the use of formulaic language will upgrade the quality of the language EAP students use (Al Hassan & Wood, 2015).

The study of formulaic language has interested linguists who study language variations. The interest stems from the belief that formulas tend to be very context-dependent; hence, they are not associated with particular communities. However, they are associated with particular genres and registers. They also indicate how texts differ and give the learners clues to understand the meanings of different texts (Durrant, 2018).

In dealing with formulaic language, there are two different perspectives in the scholarship; on the one hand, scholars like Hyland (2012) have looked upon formulas to "reveal lexico-grammatical community – authorized ways of making meaning" (p. 135). On the other hand, some scholars consider formulas to be "triggers to help thinking" (Davis & Morley, 2015, p. 28). Formulas also function as scaffolding by offering ready-made chunks of language that help learners construct academic texts. Even though they help students form language structure, they also limit and restrict originality in their writing, which is important to consider when investigating the impact of formulas on second-language acquisition (Durrant, 2018).

2.3 The Relationship Between Formulaic Language and Students' Progress in Academic Writing

Language researchers have investigated the relationship between mastering the formulaic language and overall improvement in language proficiency by focusing on two types of learners' proficiency: the sequences that frequently appear in the formulaic language of native speakers, and the types of sequences that are frequent in the learner's corpora. A later section discusses the impact of learners' corpora in advancing the study of formulaic language (Durrant, 2018).

Researchers have used several language theories to investigate the frequencies that appear in the formulaic language of native speakers. For instance,

Al Hassan and Wood (2015) focused on formulas taught to students to improve their IELTS writing scores. They discovered a strong correlation between the number of formulas used and the grades awarded for the writing tests. While studying the frequencies that appear in the formulaic language of native speakers, linguists like Granger and Bestgren (2014) also attempted to aggregate all the bigrams (two-word sequences) in the corpus of L2 English essays that were graded based on the Common European Framework Reference (CEFR). They examined them in the British National Corpus (BNC) and aggregated the formulas using two methods: a T-score to measure the degree of certainty that the bigrams occurred frequently, particularly for high-frequency formulas, like *other hand* and *long time*, and mutual information to measure the strength between the association of the biogram, or how often a word in a pair could exist without the other word, like in *pop music* or *vicious circle* (Durrant, 2018). The study discovered that the students who scored better used frequencies higher in mutual information and lower in T-scores. Bestgen and Granger (2014) also conducted a study that used the corpus of contemporary American English and found the same results: the mutual information scores of bigrams correlated with higher scores and higher quality of text which eventually led to better grades.

To investigate the impact of formulaic language on upgrading students' quality of writing, many linguists have studied the occurrences of formulaic language in learners' corpora. However, their focus was not on finding formulaic language in the academic community. They wanted to find the occurrence of formulaic language in the students' academic writing (Appel & Wood, 2016; Biber & Gray, 2013; Chen & Baker, 2016; Durrant, 2018; Staples et al., 2013; Vida-Kovic & Baker, 2010).

Since the present study was conducted in a bridge program for students who still needed to sit for a standardized tests to start their undergraduate program, it is important to consider some past studies that investigated the occurrence of formulaic language in similar contexts. As mentioned, students in the ENGB have to take some standardized tests to determine their language proficiency, such as the test for the Teaching of English as a Foreign Language (TOEFL). Similarly, Durrant (2018) introduced three studies that examined the formulaic language in the writing part of standardized tests. I will dedicate a whole

section to explain what kind of formulaic language was included in this study, as formulaic language is a broad term and includes a lot of language formulas.

The studies discussed below are part of linguists' ongoing attempts to examine formulaic language's effect on upgrading the writing quality in EAP classes.

Staples et al. (2013) investigated the use of lexical bundles in different levels of TOEFL writing tests. Surprisingly, they found that the higher the language proficiency, the less dependency on the lexical bundles. They concluded that students with lower language proficiency tended to use more lexical bundles in their TOEFL writing. These findings fall in line with those of Appel and Wood (2016), who compared the TOEFL writing of high-achiever and low-achiever students in Canadian Academic English Language (CAEL) programs and discovered that low achievers used a significant quantity of lexical bundles adopted from the prompts of the texts (Durrant, 2018). They tended to do this because, as Biber and Gray (2013) argued, the lower-level students needed to develop fixed expressions of their own or needed more language expressions to help them express themselves, whether in writing or speech. For that reason, they tended to overuse the lexical bundles and collocations (Durrant, 2018).

These findings illustrate that there is more to understand regarding the impact on learning of lexical bundles and collocation in improving the written skills of ESL learners. Linguists need to uncover how learners look upon lexical bundles, whether or not they consider them a learning tool.

2.3.1 Explicit Teaching of Collocations and Lexical Bundles

The study aims to explore the correlation between the teaching of collocations and lexical bundles and the improvement of the writing skills of first-year students. It is essential to understand the research questions better to discuss some of the early studies that implemented an explicit teaching pedagogy of the formulaic sequences to examine how such an approach contributes to a better understanding of the teaching materials.

One early intervention study that investigated the impact of the explicit teaching of the formulaic sequences was conducted by Boers, Eychmans, Kappel, Stengers, and Demecheleer (2006). The study was influenced by the "text

Chunking “introduced by Lewis (1997). The study was conducted on EFL learners in a school year course. The participants were divided into two groups; one group was explicitly taught the formulaic sequences they encountered in the reading and listening texts, and this group was identified as the text-chunking group. The other group was identified as the comparison group, and their classes did not explicitly teach them any of the formulaic sequences they encountered in the text. After the intervention, an oral proficiency interview was conducted. The text-chunking group was more prepared and used more formulaic sequences than the comparison group; however, the more significant number of formulaic sequences that the text-chunking group used did not confirm that it was the result of the uptake from the course. It is argued that the intervention helped the text chunking group to have some strategic advantage because they encountered the formulaic sequences earlier in the course. However, there is no evidence to what extent the text-chunking activities helped the students increase their knowledge of the formulaic sequences or confirm that they can carry such knowledge with them when they conduct real-life conversations outside the classroom.

The demand to investigate the impact of explicit teaching of formulaic sequences urged other researchers to conduct similar studies. Peter (2012) conducted an experimental study by using a post-test after an intervention. The participants of this study were a group of L2 German students. This study included an attention-drawing technique where the formulaic target sequences were typographically highlighted and glossed in the text. The post-test results concluded that the group with the typographically highlighted and glossed formulaic sequences performed better than the comparison group, which adds to the explicit evidence that teaching formulaic sequences enhances the student intake of the formulaic sequences.

More studies were conducted to investigate the impact of "output "activities, which require the participants of studies to reuse the formulaic target language taught via "input" activities during the intervention study. Szudarski (2012) conducted a study investigating the impact of learning verb-noun collocations on intermediate students reading proficiency. The participants of the study were assigned some reading texts. They were instructed to use one of the three conditions: read only the text, finish output activities that focus on the verb-noun

collocations in the reading text, and last, a control condition that is not exposed to any exercises. The test results at the end of the intervention indicated that the output exercises helped the participants acquire the verb-noun collocational knowledge better because they were exposed to both receptive and productive activities that led to better learning.

The efficiency of the explicit instruction of the Formulaic language has been proven successful in a study that was conducted in Saudi Arabia on a group of pre-intermediate learners in an English language program at one of the universities. The study was conducted on 81 students to investigate the effectiveness of instruction of Formulaic language in prewriting vocabulary activities and how such instructions will impact their writing skills at the end of the program. The study took ten weeks and was an experimental design study with a pre-test and post-test. The results of the study verified that the students benefitted from the taught formulaic language, and with the help of the formulaic language, they could express themselves successfully in well-developed sentences (El-Dakhs et al,2017).

Despite the increasing number of studies investigating the impact of formulaic language teaching on language proficiency, very few studies discussed in detail how formulaic language is represented in educational materials and textbooks (Pellicer-Sánchez & Boers, 2018). I will discuss this idea in the recommendation section; however, it is essential to present some of the researchers' contributions in the matter. One of the earliest studies that addressed the need to develop material to teach Formulaic language was Boer's, Deemecheleer, Coxhead, and Webb's (2014). The study introduced some of the activities used in the EFL classes to teach formulaic language; the exercises mainly fill in the gaps in which every collocation is brought down to constituents, and the students must reassemble them. The exercises all focused on trial and error. Boers et al. (2014) reported that the students' gains from such fill in the gaps exercises were surprisingly poor, and the post-test results indicated improvement by only 5-10 % more than the pre-test.

Pellicer-Sánchez & Boers (2018) argued that many learning materials in the textbook that teach formulaic language focus on learning via trial and error. Boers et al. (2017) analyzed over 323 exercises taken from 10 different textbooks. They

concluded that most of these exercises used trial and error strategies in their activities, especially when teaching collocations. They never provide the users with samples or examples of the collocations, and many of the exercises depend on guesswork. Besides, the textbook relies heavily on the teacher's corrective feedback to explain the wrong choices of the book users. The data from the classroom experiments of both Boer et al. (2014) and (2017) indicated that the wrong choices of the users might leave undesirable traces in the learners' memory despite the teachers' corrective feedback. The studies suggested that the only way to avoid such a mistake is to provide a list of the collocations prior to the exercises, and it should not be left only to the teachers' corrective feedback. The users of the textbook should have the list of collocations available to them.

The previous section introduced some studies that had explicit teaching of formulaic language as their pedagogy—having discussed the benefits of the explicit teaching of formulaic language and supported it by noticeable gains of learning. It is advisable to approach explicit teaching with care and have a realistic attitude toward it. Explicit teaching formulaic language does not mean copying and memorizing the formulaic language; on the contrary, it means constantly interacting with the target language through noticing, retrieving, and generating the target language to learn and use it. I will discuss in detail in section 3.7 how the intervention study implemented the explicit teaching of collocations and lexical bundles to improve the writing skill of First-year university students. The effectiveness of any pedagogical procedure depends mainly on how the teacher designs the activities, the amount of feedback the learners receive, and how they can build on such feedback to improve their learning (Pellicer-Sánchez & Boers, 2018).

Before defining what, the study means by collocations & lexical bundles in section 2.5, It is essential to discuss in the next section the shift of focus from single-word knowledge to multi-word knowledge and the role that corpus linguistics has played in facilitating this shift.

2.4 Learner Corpora and the Shift from Single-Word to Multi-Word Knowledge

The shift of focus to multi-word knowledge was due to the increasing role of corpus linguistics in drawing attention to pre-patterned word combinations. New

technologies have helped analyze the corpora and sample the L2 varieties, referred to as “learner corpora” (Gilquin et al., 2007).

Advancements in learner corpora studies have paved the way for more research on the impact of formulaic language on second-language acquisition. Learner corpora are of great benefit because they contain a large number of single words and multi-words, which are helpful for any study in the field of second-language acquisition. Fewer constraints on L2 learners’ choices is believed to be helpful when investigating formulaic language (Granger, 2018).

Despite this, there are constraints to using learner’s corpora in investigating language acquisitions, as corpora can only be used to examine the formation of formulaic language and not to explain how the learner perceives or learns those formulas. Learner corpora also only present learners’ performance data, which is considered an imperfect mirroring of learners’ competence; hence, language researchers argue that learner corpora should complement experimental studies and not substitute them (Gilquin & Gries, 2009).

There are many types of language structures that follow under the umbrella of formulaic language. but just two types of formulaic language are most evident learners’ corpora studies. The two types are collocations and lexical bundles.

Granger (2018) analyzed approximately 50 learner-based studies and found static versus developmental perspectives across the studies and differences in the target language and medium (Siyanova-Chanturia & Pellicer-Sanchez, 2018). First, Granger (2018) found out that most studies investigating the use of formulaic language focused on effect of formulaic language at a particular point of time and always compared the learners’ data to that of native speakers. This is exemplified in Nesselhauf’s (2005) study, which examined the use of collocation by learners of German mother tongue (L1), and also in Chen and Baker’s (2016) study, which examined the use of the lexical bundle on Chinese students who were studying English as a foreign language. Although these studies were all static, there has been a growing interest in developmental studies in recent years to examine the process of L2 learners over a longer period and in more longitudinal studies to test the progress of the students of various proficiency levels (Gass, 2013, cited in Siyanova-Chanturia & Pellicer-Sanchez, 2018).

Second, by examining the medium that most of the studies target, it is clear that majority of the studies that involved learners' corpora investigated the use of formulaic language in writing more than speech. It is imperative to include more studies of formulaic language in speech because the comparison between written and spoken data draws attention to the difference between the quantity and the quality of formulaic language in the two mediums of language. Such diversity in research will lead to a better understanding of the impact of learning formulaic language as learner corpus-based studies on writing examine essays, whereas learner corpus-based studies on informal speech examine informal interviews (Siyanova-Chanturia & Pellicer-Sanchez, 2018).

2.5 Collocations and Lexical Bundles as Two Leading Representatives of Formulaic Language

To investigate the correlation between the explicit teaching of collocations & lexical bundles and the improvement of the writing skills of First- year university students, It is imperative to define the collocations and lexical bundles clearly and explain why they are considered two essential representatives of the formulaic language because, as previously argued by Wray (2002), formulaic language is an umbrella term. The term "formulaic language" has many features, such as idioms, phrasal verbs, collocation, and lexical bundles. Using such an umbrella term is risky because it hides the diversity of the phenomenon discussed. Suppose we continue to address formulaic language as one entity. In that case, we will deprive English language learners of a good learning opportunity because every constituent of formulaic language represents different challenges for the learners. The section will discuss the unique characteristics of both collocations & lexical bundles and the significant role they can play in improving the writing skills of foreign language learners if they are implemented accurately in the English for Academic Purposes programs. Granger (2018) conducted extensive research on the different features of collocations and lexical bundles, dividing both into the components of definition and operationalization, frequency, accuracy and appropriacy, L1 transfer, and development. The following sub-sections expand on these components.

2.5.1 Components of Collocations

2.5.1.1 Definition and Operationalization of Collocations

According to a frequency-based approach to phraseology, initiated by Sinclair (1987), *collocation* refers to lexical items that occur within a certain distance and with higher frequency. Collocation is vital in producing more advanced sentence structure, yet it is challenging to master for L2 language learners. Collocation is perceived as a continuum governed by some degree of restrictions. This can vary from little restriction, like in the traditional collocation of *write an essay*, to a more restrictive collocation or frequency-based collocation, like *conduct research* instead of *make research*. This can then reach the frozen expression collocation, like *generally speaking.* The distinction between traditional collocations and frequency-based collocations is challenging to extricate. This is what Hoey (2005) proposed as "lexical priming," arguing that collocations offer clues to language learners on how they should construct sentences (Paquot & Granger, 2012). However, Nasshehuf (2005) addressed the issue of the overuse and the misuse of collocation as "collocational teddy bears" (p. 69) because learners feel more confident as they have ready-made chunks of language that can help them communicate their ideas.

On the question of measuring the associations between collocations, linguists suggest that, in the case of the frequency collocations, which occur very close to each other and in a more systematic manner that is greater than chance, statistical tests are best, such as the mutual information (MI) or T-score tests (Biber et al., 1998, cited in Granger, 2018). Thus, Granger (2018) suggested that collocations formed because of frequency are called "statistical collocation," whereas collocations formed based on a traditional formation are called "restricted collocations" (Cowie, 1998, p. 6).

Although it is natural for native speakers to recognize and understand collocations, it is challenging for non-native speakers to acquire and produce them successfully (Acherman & Chen, 2013). Nation (2001) argued that collocations are challenging for non-native speakers because they carry some grammatical and lexical unpredictability within their meaning (Acherman & Chen, 2013).

2.5.1.2 Frequency, Accuracy, and Appropriacy of Collocations

In their continuous attempts to understand collocations, researchers questioned the types of collocations and how they are formed. Tsai (2015) suggested making a distinction between “collocation density” and “collocation diversity” (p. 728, cited in Granger, 2018). Based on such distinctions, researchers identified an overuse in collocation density to the extent that some argue that language learners tend to produce fewer collocations than native speakers (Laufer & Waldmen, 2011). By looking more closely at the density of collocations, Vincent et al. (2016) also noticed an overuse in verb + noun (*V+N*) collocations and an underuse in adjective + noun (*Adj+N*) collocations, but more investigation is needed into how collocations are acquired and used among learners (Granger, 2008).

As for collocation diversity, studies on learners’ use of restricted or statistical collocations have found a lack of diversity of collocations in learners’ use of language. To investigate the use of collocation diversity, Tsai (2015) conducted a study on the statistical verb-noun collocations of Taiwanese learners of English and found a high density of collocation use but limited diversity in the use of collocations. Tsai (2015) explained that learners need to cling to the collocations they are familiar with. This goes in line with Wang (2016), who proposed that learners do not exhibit diversity in their use of collocations after looking at the verb-noun collocations in the BNC and noting that learners depend on frequent combinations, such as *make+decision* or fixed combinations like *give+rise to* and *make+use to* (Granger, 2018).

When investigating learners’ use of collocations, it is essential to consider if the learners use the acquired collocations, whether restricted or statistical. Wang and Shaw (2008) suggested that the learners’ mother tongue plays a significant role in whether they appropriately use the collocation. For example, in a study on Swedish learners, Wang and Shaw (2008) found that, due to the closeness of Swedish to English, the learners made some bold decisions in their language use and were not hesitant to make mistakes. Conversely, Chinese students, whose language is entirely different from English, were very reluctant to produce English language because they were worried about making mistakes (Granger, 2018).

Although these studies reveal how language learners use collocation, they fail to explain one important aspect: the criteria for assessing the accuracy and appropriacy of collocation use. Any language researcher needs to understand the criteria these studies applied to test the accuracy of collocations. To avoid such weakness in research, Durrant and Schmitt (2009) analyzed learners' use of collocations using a scale to measure the strength of the collocations. This method helps shed light on how the learners produce collocation because understanding how learners develop proficiency in using collocation is far more important than merely reporting that collocations are used correctly or incorrectly (Granger, 2018).

2.5.1.3 Transfer from Mother Tongue and the Development of Collocations

The mother tongue, or L1, plays a significant role in how learners receive, process, and later acquire the target language. Learner corpora studies indicate that any language learner has a reservoir of formulaic sequences in their lexicon, and this reservoir profoundly impacts how they acquire the formulaic sequence in L2. Many factors affect the transfer of formulaic sequences from L1 to L2, including the compatibility of the collocations and whether the introduced formulaic language has a similar one in the learners' L1 reservoir of formulaic sequences. In a study on French learners, Gilquin (2007) found that most of the collocations the learners used were compatible with the French ones, such as *make=faire un effort*.

Many studies that investigated the impact of L1 on learners' choices of collocations reported a high percentage of L1 impact on learners' errors; for example, Nesslehauf (2005) reported a 48% impact of L1.

The problem is not only in the incompatibility between the colocations of L1 and L2. In most cases, the problem relies upon how learners misuse the compatible collocations, which is commonly known as "L1 avoidance," as learners make mistakes in using the compatible collocations because they are afraid to make mistakes (Alonso Ramos et al., 2010).

It is important to examine the research on collocations and to investigate some of the gaps in the research. When evaluating the extensive studies on the use of collocations in improving language proficiency, there are two significant essential findings. First, developing interlanguage mastery takes time. Using a pseudo-longitudinal approach, Laufer and Waldman (2011) concluded that there is

a relationship between learners' proficiency levels and their use of restricted collocations. In a similar study, Durrant and Schmitt (2009) urged that learners tend to overuse high-frequency collocations such as *hard work*, characterized by high T-score and underuse lower frequency (Granger, 2018). By adopting Durrant and Schmitt's (2009) approach, Paquot (2017) argued that the mean MI significantly increased from level B2 (higher intermediate) in Common European Framework Reference (CEFR) to C1 (advanced) until it reaches C2 (very advanced).

Another finding is that collocational acquisition requires some time to develop. Any researcher who aspires to investigate the role of collocations in language proficiency needs to design longitudinal studies, as the results of short studies will not give accurate or reliable results. This is shown in the results of the study by Yoon (2016), which aimed to investigate the verb-noun collocational in argumentative and narrative essays by ESL learners. The study endured one semester and led to no significant development in the MI scores on any of the essays.

Similarly, Bestgen and Granger (2014) adopted the same approach of investigating a collocational study in a short period. They came up to a similar conclusion that there was no significant change in the MI score; in fact, they noted a decrease in the use of collocations by the learners. These findings reveal that studies on the acquisition of collocation need to be conducted over a more extended time and probably repeated more than once because the acquisition of collocation does not come overnight.

2.5.2 Components of Lexical Bundles

2.5.2.1 Definition and Operationalization of Lexical Bundles

Studies on learner corpora have taken a new direction by introducing lexical bundles, which are considered an asset in the field of phraseology (Granger, 2018). The lexical bundle is a term adopted by Biber et al. (1999), and it refers to high-frequency combinations of words that often occur next to each other, whether in spoken or written text. Lexical bundles are incomplete grammatical structures, commonly formed by combining two noun phrases or a verb and noun phrase or more phrases or clauses, such as *on the other hand* and *you are responsible for* (Durrant, 2017). Lexical bundles are considered building blocks of any spoken

or written text; despite their unique nature as extracted chunks of language, they cannot deploy any syntactic integrity because they cannot stand on their own and need to be incorporated into sentences (Ädel & Erman, 2012). Lexical bundles have three advantages that interest language researchers: they can be easily identified in any given text, they play functional roles, and they can demonstrate the differences between text types (Durrant, 2017).

Before discussing the operationalization of lexical bundles, it is essential to consider the different types that would help improve the writing skills of EFL learners. The study focused on introducing the lexical bundles that would help the study participants improve their writing skills. The study introduced different types of lexical bundles weekly throughout the 15-week programme. The study aimed to train the participants to develop well-organized essays at the end of the programme. To reach this aim, I designed a detailed weekly course to introduce the different collocations and lexical bundles to help the participants logically develop their writing.

This section is dedicated to explaining some of the essential key terms that were mentioned in Table 2. To begin with, the course classified the lexical bundles in this study based on the classification that was introduced by Hyland(2008), Biber & Barbiere(2007), Biber, Conrad and Cortes (2004), Cortes (2004) and Simpson-Vlach and Ellis (2010)

Those prominent scholars divided the lexical bundles into three categories: referential bundles, discourse organizing bundles, and attitudinal bundles. Under every category, there are subcategories. First, the referential bundles, which are also considered research-oriented bundles that represent how reality or ideas are represented; for example :

Time/ text bundles – *at the end of the*, *at the beginning of*, attribute bundles – *a little bit of, the use of*, and lastly, the topic-specific bundles – *in the interpretation of*. Secondly, the discourse organizers bundles are text-oriented bundles in Hyland's (2008) list. Those bundles are mainly concerned with how the writers introduce their ideas, develop them through writing and eventually build a solid argument that reflects a well-developed line of thought; for instance: logical relations bundles are also referred to as transition bundles – *on the other hand*, *in contrast*. Finally, the attitudinal bundles are also stance bundles in Hyland's (2008) list. Hyland (2008)

argued that those bundles could be referred to as participant-oriented bundles because they demonstrate interpersonal meaning; for example, the interactional bundles are used to argue a point – it *should be noted that* – it can be seen (Dontcheva-Navratilova,2012).

To discuss lexical bundles' operationalization, we need to consider bundle size, frequency, and dispersion. Language researchers vary in their views on lexical bundle size. Some argue that they could be bigrams (two-word sequences; Crossley & Salsbury, 2011), but the bundle size could be three words (Paquot, 2013), four words (Chen & Baker, 2016), even up to six-word bundles (DeCock, 2000, 2004).

There is a relationship between the number of words and the frequency, as well as the occurrence of the bundle per million words. As for the dispersion criteria, findings of any lexical bundles represent the corpus as a whole. Unfortunately, there is a lack of this criterion in learner corpora studies because it means that the number of occurrences of certain bundles in the writing samples of some learners should demonstrate that the learners used the lexical bundles (Granger, 2018).

2.5.2.2 Frequency, Accuracy, and Appropriacy of Lexical Bundles

It is surprising that when language researchers measure the frequency of lexical bundles across novice and expert writing, they cannot generalize the results, partly because of the adoption of the various criteria of frequency, size, and dispersion, but mainly due to the different degrees of processing after the lexical bundles are extracted (Granger, 2018). There are several reasons for this. First, lexical bundles that overlap, such as *due to the fact* and *the fact that*, can either be considered two different lexical bundles, or as one (Ädel & Erman, 2012). Second, some lexical bundles are already found in the prompt of the writing task; learners usually tend to copy them again when they write their essays. Researchers must agree whether to count them as part of the learners' lexical bundles or exclude them from the analysis (Staples et al., 2013) Third, researchers must decide whether to include all of the learners' lexical bundles, especially if some of the lexical bundles are part of target like lists that the learners have previously learned.

When discussing the frequency of lexical bundles, researchers face the same problem as with collocation: learners tend to overuse bundle tokens associated with underuse in bundle types. There is growing interest in how learners choose their lexical bundles. Researchers have found a clear overuse or underuse of some lexical bundle. For example, Ädel and Erman (2012) examined samples of Swedish learners' academic writing and noted an underuse of "this" bundles (e.g., *this may be because*), which are considered a significant feature of academic writing. Another primary observation is some overused lexical bundles do not match the register. Learners tend to transfer bundles that belong to the spoken register in their academic writing (Granger, 2018).

Studies often compare the frequency of lexical bundles to learners' native corpora, disregarding the degrees of association that hold the words of the lexical bundle. Thus, researchers have measured the lexical bundles based on statistical association scores (O'Donnell et al., 2013).

The question of the accuracy and appropriacy of the use of the lexical bundle is debated as language researchers do not consider lexical bundles that are not part of the native corpus. However, L2 EFL learners tend to use some lexical bundles that belong to them and that might not exist appropriately in the native corpus. Such lexical bundles are what De Cock (2000) refers to as "idiosyncratic learner combination[s]" (p. 58).

There are many examples of lexical bundles that frequently occur in L2 EFL writing and that do not necessarily have a meaning in the native language; *on the other side* is often used as a connector to mean *on the other hand*. Another example is the excessive use of *according to me* instead of *in my view* or *in my opinion*. The analysis of L2 writing demonstrates that students use some bundles that do not exist in the native corpus, although those lexical bundles do not have correspondence in the native corpus or do not occur enough to reach the frequency level (Granger, 2018).

The problem of appropriacy lies in the issue of the mismatch of the register. L2 learners often use lexical bundles that belong to the spoken register in their academic writing, and they also underuse some of the academic lexical bundles that fit the academic register. For instance, Granger (2017) noticed an overuse of verb-based bundles (for instance, *we can say that*), which belong to the spoken

register, inserted in essays. There is also an underuse of lexical bundles like *in the case of*, which should be evident in any academic essay (De Cock, 2000, p. 58). The mismatch is not always in the register; it could be semantic or pragmatic. Chen and Baker (2016) argued that learners often semantically misuse *on the other hand* because they employ it as a multifunctional connector between all types of clauses, and not necessarily clauses that contain contrasting ideas.

2.5.2.3 Transfer from Mother Tongue and the Development of Lexical Bundles

Mother tongue transfer has never been the focus of studies on lexical bundles. The focus has been on the existence of similar lexical bundles in the learners' L1 that might affect the production of the L2 lexical bundles (Granger, 2018). When discussing the development of the lexical bundles L2 learners use, most studies adopt a pseudo-longitudinal approach in an attempt to compare the quantity and the quality of lexical bundles on the different proficiency levels of L2 learners. Staples et al. (2013) concluded, after investigating the use of lexical bundles across the three proficiency levels in a TOEFL course, that the higher the proficiency level, the lower the dependency on lexical bundles. In a similar longitudinal study, Crossley and Salsbury (2011) investigated the accuracy of the use of bigrams, or two-word bundles, and concluded that the use of bigrams improved over the year. The common factor in the research on both collocations and lexical bundles is the timeframe. Researchers insist that anyone interested in investigating these two features of formulaic language should conduct longitudinal studies to reach some overall results that can represent solutions to the proposed research questions.

2.6 Challenges of Developing Testing of Formulaic Language Among L2 Learners

Formulaic language has significantly impacted English-language standard teaching practices. The term *formulaic language* has a number of different language categories under its umbrella, as discussed above. One of the most significant challenges of formulaic language is the difficulty of developing appropriate exams to test formulaic language knowledge. Pawley and Syder (1983) noted that the number of formulaic sequences that make up formulaic

language can be several hundreds of thousands. As such, the lexicon of formulaic language is way larger than single-word vocabulary (Jackendoff, 1995). It is challenging to develop tests of formulaic language due to the difficulty of identifying all its target language uses and the massive number of the lexicon (Gyllsad & Schmitt, 2019).

It is essential to explain why the diversity of formulaic language hinders the development of tests. First, formulaic language is measured differently; some formulaic sequences are measurable because they contain statistical criteria, as explained in the MI scores in tests that involve collocations. Conversely, some aspects of formulaic language are difficult to assess accurately because their criteria are very subjective; thus, assessing the successful use of idioms depends heavily on the criteria set by each researcher (Grant & Nation, 2009). Some linguists argue that, over the years, formulaic language can be stored holistically, which makes it impossible to measure because such storage will differ from one learner to another:

> "the means of storage and retrieval of the same sequence can differ from one individual to another, and can differ from one time to another for the same individual depending on the a range of factors such as changes in proficiency, changes in processing demands, and changes in communicative purpose" (Read & Nation, 2004, p. 25).

These challenges have led to a lack of standardized testing and assessment of formulaic language; however, this has not stopped language researchers from attempting to develop tests that can examine the successful acquisition of different categories of formulaic language, such as collocations, idioms, and phrasal verbs. The subsequent sections discuss seven types of tests that demonstrate some essential factors in the progress of the testing of formulaic language (Gyllstad & Schmitt, 2019).

2.6.1 The Word Associates Test (WAT)

Read (1993) developed the Word Associates Test (WAT) to measure the academic English vocabulary of 800 wordlists based on their frequency in academic texts. The words were introduced in the University Word List (UWL). The test was later developed to test the matching collocation. The test only examines

the strong association between collocations. As seen below, the test-taker has to circle all of the possible answers that strongly associate with the target item, in this case, the adjective *sudden*. The correct answer would be to circle *quick* and *surprising*, which are considered synonyms of *sudden*. The test taker would also need to circle *change* and *noise*, which are considered collocations (Gyllstad & Schmitt, 2019).

Sudden

Beautiful	Quick	Surprising	Thirsty
Change	Doctor	Noise	School

As shown above, WAT does not test knowledge of collocations in general; it is designed to test the depth of knowledge of some target collocations. In general, the knowledge of collocations is considered to happen as a later stage of lexical knowledge (Siyanova-Chanturia, 2015). The mastery of collocation is also considered far more complicated than the receptive knowledge of vocabulary (Laufer & Waldman, 2011). Researchers have found that WAT is not an accurate testing method due to the multiple-choice design of the test that gives more room for guessing. WAT was developed over 23 years ago. It is still in the process of development due to the complexity of measuring the depth of knowledge of collocations, one of the components of formulaic language (Gyllstad & Schmitt, 2019).

2.6.2 The COLLEX, COLLOMATCH, and DISCO Tests

Gyllstad (2007) developed the COLLEX and COLLMATCH tests to measure the receptive verb-noun collocation among upper secondary and university students in Swedish EFLs. The COLLEX test consists of 50 items, and the test-taker has to choose the verb-noun combination that is common and most frequent in the English language. An example of this taken from Gyllstad (2009, p. 157) is:

A. Drive a business B. Run a business C. Lead a business

The COLLMATCH is a yes/no test composed of decontextualized verb-noun collocation, and the test-taker has to decide if the collocation is among the most frequent and commonly used ones in the English language. An example taken from Gyllstad (2007, p. 309) is:

1. Have a say yes/no
2. Lose sleep yes/no
3. Do justice yes/no
4. Draw a breath yes/no
5. Turn a reason yes/no

Gyllstad **(**2019) administrated both tests alongside the Vocabulary Language Test (VLT), which was is developed to measure single-word vocabulary (Schmitt et al., 2001). Surprisingly, the results of the COLLEX and COLLMATCH tests correlated with the VLT test. Despite the different formatting of COLLEX and COLLOMATCH, as the COLLEX consists of 50 multiple choice questions, whereas COLLOMATCH consists of 100 items in yes/no questions, they both provide similar information. A drawback of these two tests is that there is no demonstration of knowledge, as the formatting of the tests leave a chance for guessing, making the test administrator unsure whether the answers are based on actual acquired knowledge or a strike of luck (Gyllstad & Schmitt, 2019).

Another test is the Discriminating Collocations Test (DISCO), designed to measure the receptive knowledge of collocation. The test has been administrated after 60 hours of language instruction to test the learners' idiomatic knowledge by asking them to choose which two-word collocation is an idiom in English. An example from Eyckman (2009, p. 146) is:

A. Seek advice B. Pay attention C. Express charges

Eyckmans (2009) introduced DISCO in an attempt to test collocations. However, it did not give any indication of whether the learner could incorporate formulaic language in fruitful spoken conversation. Notably, none of the above tests provided valid or reliable evidence to prove how the test results could be interpreted to offer overall collocation knowledge (Gyllstad & Schmitt, 2019).

2.6.3 CONTRIX Test

Revier (2009) developed the CONTRIX test to measure the productive knowledge of verb-object noun collocation. The test-taker receives a prompt in the form of a sentence that has gaps in it, and their task is to fill in the gaps by

choosing from three choices to form the verb–object–noun collocation (e.g., *tell the truth*). An example from Revier (2009, p. 129) is:

The quickest way to win a friend's trust is to show that you are able to _____.

Tell	a/an joke	joke
Take	the	secret
Keep	----	truth

This format was adopted from Schmitt's (2010) matrix, which is used to test learners' ability to recognize different types of word knowledge. However, Revier (2009) argued that this test format could not test the productive knowledge of test-takers because they do not merely choose to combine the lexis; on the contrary, test-takers should demonstrate some sort of grammatical knowledge by encoding the noun and the verb constituents to form the verb–object–noun collocation successfully. Unfortunately, the test-takers' results across different proficiency levels did not prove that the test confirms the learners' productive knowledge of collocations (Gyllstad & Schmitt, 2019).

2.6.4 A Productive Collocation Test

Developing tests that can measure learners' productive knowledge has long been essential to formulaic language studies. Schmitt et al. 2004) developed a productive test to measure the productive knowledge of collocation. In this test, a group of previously taught collocations are embedded in a paragraph. The first letter of the collocation is omitted, and, next to the paragraph, there is the synonym of the collocation. The test-taker has to complete the collocation. The test aims to evaluate learners' knowledge the "form" of the collocation and not the "comprehension" of its meaning. An example from Schmitt et al. (2004, pp. 58–59) is:

Learning English as a second language is a difficult challenge, but we do know several ways to make more efficient. F--------- of---------, almost every research study shows that you need you use English as much as possible.

(the initial one)

I ----------is cl---------- that the more you use English, the better you will learn it. There is no disagreement about this. (this is obvious)

Notably, the format of this test was adopted from that of Laufer and Nation (1999), and it was also used in the VLT on single-word vocabulary. This test overcomes the element of guessing that was a major drawback in the previous tests. This test shows that the test-taker knows the collocation and has what Schmitt (2010) called “form recall” (p. 89). The results also show that that the test-taker knows how to spell the target collocation; hence, it gives a little evidence that they can use the target collocation successfully in the other four skills.

2.6.5 The PHRASE Test

Martinez (2011) developed the PHRASE test. The promising aspect in this test is that it is derived from a 505 phrasal verb list. The design is quite simple; the test-taker has to read a decontextualized sentence that contains a phrasal expression, and they then have to choose a synonym to the highlighted phrasal expression from four options. An example from Martinez (2011, slide 54) is:

At once: I did it at once.

a. one time
b. many times
c. early
d. immediately

The fact that the phrase test is sampled from a list of phrasal expressions is considered a step in the right direction in designing tests to measure productive language knowledge because the test analysis can interpret the results compared to overall size (Gyllstad & Schmitt, 2019).

2.7 Characteristics of Developing Formulaic Language Assessments

The previous attempts at developing formulaic assessments have proven that it is challenging to create tests that provide reliable and credible results. In their continuous attempts to investigate the impact of learning formulaic language on the language acquisition, Gyllstad and Schmitt (2019) came up with a set of

principles that can help educators develop tests to measure the productive knowledge of formulaic language.

2.7.1 Establishing a Construct

One of the drawbacks of the previous tests is that they need a reference or, as Gyllstad and Schmitt (2019) call it, a *construct*. The tests holistically assess the acquisition of formulaic language but do not specify what type of knowledge they are assessing. Bachman (1990) proposed that educators who plan to develop an assessment for a formulaic sequence, whether collocation or lexical bundles, should follow the following steps: defining the construct theoretically, defining the construct practically, and establishing comprehensible procedures for collecting data. Having completed these steps, the researcher should consider the items to include in the construct. The key aspect of establishing a construct is establishing a purpose before developing the assessment.

2.7.2 Selecting the Appropriate Formulaic Categories for Assessments

Due to the unique, diverse nature of formulaic language, it is difficult have a single standardized test that can fit all the categories of formulaic language. Selecting the right formulaic category is a crucial step in determining the purpose of the assessment. Test developers need to decide what the test is trying to assess, whether the knowledge of collocations or lexical bundles. Alternatively, it might test how learners can successfully incorporate them in their writing tasks. After determining the purpose, a clear definition of the formulaic category should be given to the learners.

There have been many attempts to aggregate vocabulary in what is called the Academic Word List (AWL). These are the most widely used list used in the field of English for Academic Purposes (EAP). However, linguists have started to question the validity of the single-word list as research is more inclined toward word co-occurrence rather than single words. Unlike AWL, the Academic Collocation List (ACL) does not focus on a single word; it focuses on lexical collocation.

The ACL is derived from the component of the Pearson International Corpus of Academic English (PIACE). The corpus contains over 37 million words of

academic writing and speech from the five major English-speaking countries. The written curricular component also contains 25.6 million words taken from journals, articles, and textbooks.

The ACL was developed over various stages; it started with a computational analysis of the written component, followed by a manual refinement of the data based on quantitative parameters and target parts of speech combinations. The result was given to a language expert to determine whether each collocation was pedagogically relevant (Acherman & Chen, 2013).

Another list similar to the ACL is the Academic Formulas List (AFL) developed by Simpson-Vlach and Ellis (2010). The AFL aims to categorize the formulas according to their function, and Simon-Vlach and Ellis (2010) provided a list of compare and contrast formulas that included some of the collocations that learners can use in their writing tasks.

2.7.3 Sampling

Sampling is a fundamental characteristic of developing any test because it indicates the knowledge of the construct. For an example, if learners can answer several questions correctly about the (-ed) past tense of a verb form, this is a good indication that this rule could be used to teach a past tense verb. Otherwise, it will be difficult to apply the same rule to teaching formulaic language. Unlike rule-based construct, different categories of formulaic language are item-based constructs, which means that if a learner knows some categories of formulaic language, this does not necessarily mean that they know the other aspects of the same formulaic language.

Gyllstad and Schmitt (2019) urged researchers to select more representatives from the overall population of the formulaic language to examine the samples to predict their overall knowledge of the formulaic language. To illustrate, Schmitt et al. (2001) wanted to sample a 3,000 level of VLT, and they chose 30 items to represent the 1,000 words in the level. If the learner answered 15 questions correctly, this was an indication that they would know 500 words or 50% of the 1,000-level word. Unfortunately, there is little research to determine whether this adopted sampling approach could work successfully with an item-based construct. Collocation and lexical bundle test developers need to run their

validation to determine the sampling rate appropriate for the tests. They must also determine the appropriate number of items that are considered sufficient in a test to assess the productive knowledge of the mentioned categories of formulaic language.

From the recommendations of Gyllstad and Schmitt (2019), it is quite evident that one of the significant challenges facing developing a more standardized EFL test is the need for more validity. Language researchers should investigate ways to develop more validated tests, and, to do so, they should follow the sequence of first establishing a construct, then determining a real purpose for the tests, providing clear instructions of the formatting of the tests, and a transparent, detailed descriptive grading scheme to explain what the test score means.

For all the above-discussed reasons, there is no doubt that collocations and lexical bundles play a crucial role in enhancing language acquisition; this is evident from the number of books and publications discussing the benefits of implementing collocations and lexical bundles in language programmes.

Despite the vast number of publications, there still needs to be more research on the types of collocations and lexical bundles that would be more beneficial in enhancing language fluency. The study is to bridge the gap in research and provide a clear explanation of the types of collocations and lexical bundles that would facilitate language learning, besides recommending the language pedagogy that would make implementing the collocations and lexical bundles more successful and beneficial in language programmes.

CHAPTER 3: Methodology, Data Collection, and Analysis

This chapter describes the theoretical framework, methodological approach, methods, and sequence adopted in the study. It also explains the adopted research design and justifies its validity and the choice of the methodological approach according to the different phases of the study. It then briefly describes the data analysis for each method and the procedures of obtaining ethical approval to conduct the study.

3.1 Theoretical Framework of the Study

This study adopted the scientific paradigm, which is used in the physical sciences and experimental psychology. The scientific paradigm is primarily concerned with objectivity and the discovery of scientific generalization that describes the subject of the study. It focuses on quantitative data, which is why it employs experimental methods. I chose to use the scientific paradigm because it aims to generalize laws, which could lead to more development in the educational process. In this study, I used the scientific post-positivism paradigm to test classroom and learner variables and to associate them with educational/learning outcomes (Ernest, 1994).

The rationale for adopting the scientific paradigm stemmed from the well-grounded belief that, in the current study, teaching and classrooms exist independently, regardless of any researcher's views on them. The present study thus examined the impact of teaching collocations and lexical bundles on improving writing quality, eventually leading to better final grades for first-year students in Western universities in the Middle East. The current study was designed to employ experimental and scientific tests to examine the effectiveness of collocations and lexical bundles. The scientific paradigm advocates that the researchers should always establish a relationship between a social phenomenon (the language classroom, in the current study) and generate a hypothesis to test through continuous observation (Grix, 2004).

Research that adopts the scientific paradigm argues that we can only explain how we know what we know via scientific explanation. According to Neuman (2003), the scientific paradigm views any social problem as a chance to

apply methods to understand the behavior through close observation, making a hypothesis, and eventually predicting a solution. In the current study context, the social problem was the incompetent writing skills of EFL freshman students. In light of the scientific paradigm, this was a chance to apply logical methods and test them through quantitative tests to examine the efficiency of the hypothesis that the teaching of collocations and lexical bundles could improve the writing outcome of a group of freshman students. I used scientific tools as tests and closely observed and analyzed the data extracted from the tests. Tolley (2004) argued that the goal of any researcher who adopts the scientific paradigm should be to remain objective in their observation of the social experiment under study to reach the closest reality.

3.2 Research Questions

The research questions guiding this study were the following:

1. To what extent do the explicit teaching of collocations & lexical bundles help students use them accurately?
2. Is there a correlation between teaching collocation and lexical bundles and improving students' overall writing grades?
3. Is there a relationship between the increase in the number of lexical bundles used and the change in essay scores after intervention?

3.3 Research Methodology

I implemented an intervention study methodology, typically used in experimental study designs with a control group and an experimental group. I applied an intervention vocabulary program that focused on teaching collocations and lexical bundles instead of the words on the Academic Word List (AWL), a common practice in the English Department at the university where the study took place. I chose to include collocations and lexical bundles in the study due to their vital role in competent English-language production; many linguists have indicated that formulaic sequences should be a significant component of classroom instruction (Li & Schmitt, 2009; Nation & Newton, 1997; Nattinger & DeCarrico, 1992; Sinclair, 1991).

Researchers who investigated the impact of teaching formulaic sequences, especially collocations and lexical bundles, on language teaching have come up with lists that can guide other teachers who plan to include collocations and lexical bundles in their classes. These include academic lexical bundles (Biber et al., 2004), phrasal expressions (Martinez & Schmitt, 2012), academic collocation (Ackermann & Chen, 2013), and phrasal verbs. Scholars have also suggested that teachers should be selective in their choice of the target formulaic sequences based on their students' levels of English acquisition (Boers & Lindstromberg, 2009; Lewis, 1997). As the current study investigated the impact of teaching the collocations and lexical bundles on the improvement of writing skills, I chose to use the Academic Formulas List, assembled by Simpson Vlach and Nick (2010), because it is divided into a written core and a spoken core. It was also validated and assembled according to a specific purpose and directed to particular learners who are very much similar to the study population (Schmitt, 2022). I chose to gather my Academic Formulas List from the written core.

The list of the collocations and lexical bundles I used in the intervention study with the control group is provided in Table 1 below. I adopted the list from Simpson Vlach and Nick (2010) and taught it in the ENGB intervention over 15 weeks. The class met four times a week for one hour. Every week, I gave the class new formulas. There was always a revision session at the beginning of every class where we recalled and revisited the previously taught collocations and lexical bundles.

The list is divided into three collocations and lexical bundles categories: referential expressions, stance expressions, and discourse organizing functions. The reason for choosing this list was that the course was mainly a writing course for first-year university students. I followed the syllabus that mainly focused on teaching students all the necessary skills to write a developed, unified, coherent essay.

Table 1

The List of Collocations and Lexical Bundles Used in the Present Study

Group A: Referential Expressions		
An attempt to	in accordance with	in the course of
Are/was based on	in such a way that	In the form of
Depend on	in the absence of	there are no
On the basis of	in terms of	there are several
With regard to	in this case of	in a number of
In terms of	degree to which	in some cases
Be related to	it has been	this does not
Which can be	does not have	this means that
Is more likely	on the other hand	the difference between
At the end of this point	at this stage	to distinguish between

Group B: Stance Expressions		
Appear(s) to be	Are likely to	As a whole
Assumed to	Be argued that	Be explained by
Be regarded as	Been shown that	If they are
Is determined by	We have seen	Take in account
Can be achieved	Most likely to	Carried out by
Has been used	It should be noted	Take into account
Can be expressed	Can be achieved	Are able to

Group C: Discourse Organizing Functions		
As shown in	Important role	It is necessary
It is obvious that	It is interesting	It is worth
It is difficult	In the present study	As a consequence,
As a result of	Due to the fact	Is affected
It follows	And if you	Even though
In conjunction	Due to the fact	To determine whether

I applied a communicative language teaching pedagogy in my study. My approach to teaching collocations and lexical bundles originated from the claim that

successful vocabulary learning has three psychological processes: noticing, retrieving, and generating (Nation, 2001). Noticing can occur when any formula, whether a collocation or a three-word bundle, is highlighted as noticeable in the reading text, from the assigned authentic reading passages in the participants' coursebook, in the input, or in a class discussion in the pre-reading task. Looking up a word in a dictionary, guessing from the text, deliberately studying the formula, or explaining the formula are all possible factors that can lead to noticing (Simpson-Vlach & Ellis, 2010). In the present study, I began my class by giving the students a checklist to test their knowledge of the formulas to be taught during the lesson (see Appendix 1). When I selected the formulas for the study, I remembered that motivation and interest are two essential factors for teaching formulas. I tried to select formulas that would benefit the students and help them later express their ideas in writing better.

The retrieving stage is when the students must understand the formula, whether collocation or lexical bundles, and understand it through explicit instruction. This process is to retrieve the taught formula, which can be receptive (i.e., recognizing the form of the word and recalling its meaning when learners encounter the word in listening or reading), or productive. I used a reading passage during the retrieval stage in my class so the students could read the formulas in context. Finally, the last stage in the process is the productive one of generation, which occurs when a student takes a word they previously encountered and reproduces it repeatedly but in a different way. Stahl and Vancil (1986, cited in Nation, 2001) advocated that teachers should encourage discussion at this stage, as discussion plays a vital role in building a semantic map, which is crucial to building good vocabulary knowledge. See Appendices 2 and 3 for some exercises I used in class to generate sentences before writing the post-test essay. I also encouraged my students to use a vocabulary card to study the different aspects of the formulas (Appendix 4).

The writing stage in the intervention study entailed practicing the writing of different types of essays by using the taught formulas. The writing stage focused on using the formulas taught to produce well-developed sentence structure essays. The second stage began by teaching the students the important features of a well-developed five-paragraph essay. The intervention study focused on how the use of

collocations and lexical bundles could develop more unified and well-supported paragraphs. Teaching the formulas went in hand with teaching the main features of writing coherent introductory, body, and concluding paragraphs. The second stage thus taught the participants how to build a strong argument using collocations and lexical bundles. The participants worked on different class activities that taught them how to use collocations and lexical bundles to support their ideas and write more coherent paragraphs that would lead to a better argument. The second stage also trained the students to use collocations and lexical bundles to write well-developed outlines that would turn into their first and final drafts.

The Intervention Scheme

Week 1		
	Topics and Readings	**Assignments**
	Course Introduction, syllabus review, policies, and procedures Essential processes of For Academic Writing Developing an Academic Vocabulary Use context Clues for General Vocabulary	
Week 2		
	Topics and Readings	**Assignments**
	Introduction to a Five Paragraph Essay Writing Effective Introductory Paragraph Create unity and Coherence in paragraphs Understand the role of the introductory paragraphs Introductory sentences and reversal transitions **What is collocation?** **What is Lexical Bundle?**	
Week 3		
	Topics and Readings	**Assignments**
	Introduction to Essay Development Important considerations in Essay Development Patterns of Essay Development The Function of Supporting details Types of Supporting Details **Strong & weak collocations** Handout of Essay 1 Exemplification Essay In class Draft of Essay 1 Diagnostic Practice Growth Quiz	

Table 2, continued

Week 4		
	Topics and Readings	**Assignments**
	Thinking Critically about ideas The different elements of critical thinking Key words and supporting Details Transitional Clues to Major Details **Grammatical Categories of collocation** Diagnostic Practice Growth Quiz	
Week 5		
	Topics and Readings	**Assignments**
	Recognizing Patterns of organization Typical Topic-sentences Multiple-Definition Paragraphs Intensifying adverbs Softening adverbs Submission of Essay 1 Diagnostic Practice Growth Quiz	
Week 6		
	Topics and Readings	**Assignments**
	The Style and Tone in Writing Writing Arguments Understanding, Outlining and Synthesizing Longer Readings **Referential bundles** **Practice of Text- oriented lexical bundles** Diagnostic Practice Growth Quiz	

Table 2, continued

Week 7		
	Topics and Readings	**Assignments**
	Understanding structure and writing Correct Sentences The Role of Inferences in Comprehension and Critical Reading Discourse organizers Handout of Essay 2 Assignment sheet Diagnostic Practice Growth Quiz	
Week 8		
	Topics and Readings	**Assignments**
	Strategies for Argument Use Tactful, Courteous Language Point out Common Grounds Acknowledge different Viewpoints Rebut Different viewpoints **Logical relations bundles** **Practice of Research oriented lexical bundles** Feedback on the Draft of Essay 2 Diagnostic Practice Growth Quiz	
Week 9		
	Topics and Readings	**Assignments**
	Using Punctuation and Mechanics Identifying Chain of references in Reading The difference between Logical and illogical Inferences **Transition & Resultative bundles** Diagnostic Practice Growth Quiz	

Table 2, continued

Week 10		
	Topics and Readings	**Assignments**
	Drawing Inferences about Supporting Details Making Connections Between Paragraphs in longer Essays Drawing Logical Conclusions	
	Interactional Bundles	
	Diagnostic Practice Growth Quiz	
Week 11		
	Topics and Readings	**Assignments**
	Understanding and Evaluating Argument Flawed Arguments Irrelevant Reasons **Stance Bundles**	
	Submission of Essay 2	
	Diagnostic Practice Growth Quiz	
Week 12		
	Topics and Readings	**Assignments**
	Developing a Reader Response Essay Considering Purpose and Audience **Practice of Participant oriented lexical bundles**	
	Handout of Essay 3 Assignment sheet	
	Phrases and Dependent clauses Diagnostic Phrases and Dependent clauses Practice Phrases and Dependent clauses Growth Quiz	

Table 2, continued

Week 13		
	Topics and Readings	**Assignments**
	Understanding the Difference Between Informative Writing and Persuasive writing Feedback on the draft of Essay 3 Parallel Structure Diagnostic Parallel Structure Practice Parallel Structure Growth Quiz	
Week 14		
	Topics and Readings	**Assignments**
	Understanding the tone in Persuasive Writing Restrictive and Nonrestrictive Clauses Diagnostic Restrictive and Nonrestrictive Clauses Practice Restrictive and Nonrestrictive Clauses Growth Quiz	
Week 15		
	Submission of Essay 3 Wrap up of the programme	**Assignments**

3.4 Methods

To analyze the results, I used a pre-test and post-test: a multiple-choice and fill-in-the-blanks pre-test and post-test, and a writing pre-test and post-test. The rationale in this was to align with the students' input. As mentioned, the teaching method in the intervention program was based on following the process of noticing, retrieving, and generating the collocations and lexical bundles in an attempt to investigate if this would improve the writing skills of EFL students and lead to improvements in their overall grade. Thus, I designed the assessment plan to align with the input of knowledge. The fill-in-the-blank and Multiple choice questions questions aimed to assess if the student could retrieve the taught collocations and lexical bundles they had encountered in previous classes (Appendix 6). In the writing tests, the students then wrote an essay in which they demonstrated their ability to generate some of the formulas in a written context. The rationale was to

assess the student's ability to incorporate the taught formulas in a morphologically and grammatically correct way in a written context to improve their overall writing grade (see Appendix 7).

3.5 Instruments of the Study

The study used two types of pre-tests and post-tests: a multiple-choice and fill-in-the-blanks pre-test and post-test, and a writing pre-test and post-test. The rationale was that each instrument would help answer a different research question guiding this study.

3.6.1 Multiple choices and fill-in-the-blanks Tests

The Pre and Post tests were designed to answer the first research question, which was concerned with what extent the explicit teaching of collocations of lexical bundles could help first-year students use them accurately in sentences. As mentioned, the rationale behind testing the accuracy of the formulas was to create a baseline and starting point for the teaching of collocations and lexical bundles. It was imperative to test whether the participants could understand the different functions of the collocations and lexical bundles taught accurately, meaning that they knew how to spell them and place them correctly in the sentence. The test forms were similar to the exercises I trained the students on in class, but the post-test included all the taught forms (see Appendix 1).

Chapter 2 offered the literature review that thoroughly discussed all the tests developed to test the knowledge of the collocations and lexical bundles. Unfortunately, the tests are still a work in progress. Schmitt (2022) argued that even 23 years after their initial development, there is no accurate data that such tests can measure the collocation knowledge of the EFL learners. However, I adopted the pre-test and post-tests from the CONTRIX test developed by Revier (2009); the test measures the productive knowledge of verb-object noun collocation. The test-taker receives a prompt in the form of a sentence with gaps in it, and their task is to fill in the gaps by choosing the correct collocations and lexical bundle (Please refer to Chapter 2.6.3 for the complete account of the test). I decided to develop the pre-and post-test tests in multiple-choice and fill-in-the-blanks formats to test the impact of the explicit teaching of the selected

collocations and lexical bundle on the students' writing skills development (see Appendix 6).

3.6.2 Writing Assessments

The writing assessments were the essays the participants wrote pre – and post-intervention study. The writing assessments were designed to answer the second and third research questions. The second research question concerned the correlation between the use of collocations and lexical bundles and the improvement in the overall writing grades of the participants. The third question investigates the relationship between the increase in the number of collocations &lexical bundles and the overall grade improvement. The study identified the most used collocations and lexical bundles and what positively impacted the grade's improvement.

The post-intervention writing assessments were scheduled later, in week 16 of the intervention. The rationale of the delayed writing assessment was to ensure that the participants comprehensively studied the collocations & lexical bundles (Appendix 1). Such knowledge was tested in the post-test, composed of multiple choice and fill-in-the-blanks. The writing assessments required the participants to write an essay consisting of an introduction, three body paragraphs, and a conclusion. The word limit should be between 750 and 850 words. The participants should choose one of the three genres: exemplification, argumentative, or cause and effect. The three genres are decided by the head of the department and part of their syllabus.

The successful completion of the course requires the participants to learn about the three specific writing genres. It is worth mentioning that the study took place in the English Language Bridge program at one of the American universities in the Middle East. All the participants were required to fulfill the learning outcomes of the syllabus. The writing assessments were to test the student's knowledge of collocations and lexical bundles and to assess their capability to demonstrate such knowledge by writing a well-developed essay. To assess the participants' writing skills before and after the intervention study, they wrote pre- and post-five-paragraph essays throughout the intervention study.

In analyzing the data from the writing assessments, I looked for the correlation between using collocations and lexical bundles and improving the overall grade. I identified the highly used collocations and lexical bundles to investigate whether the increase in the numbers of the collocations and lexical bundles had contributed to improving the overall grade. Please refer to (Appendix 7) to see the prompts and the outline of the writing assessments.

3.6 Procedures of the Intervention Study

The study took place at an American University in the Middle East. The participants met in class four hours a week for 15 weeks.

The coordinator of the English division usually contacted the whole English faculty to form the semester's schedule before the semester began. Every English faculty member committed to teaching the whole semester. The students had to attend the classes to pass their English 100 level. It was a pass/fail course with no credit but a prerequisite for ENG 101, the first credited course in English.

The study took place in a well-structured programme. All the participants should fulfill the course's learning outcomes to pass the course and move to the next level in the English courses. All the ENGB students should acquire and demonstrate the fundamentals of good writing skills from unity, support, transition, and coherence. They should be able to produce a well-structured five-body essay at the end of the 15-week semester. The intervention study had to be planned around the objectives mentioned in the syllabus. The below syllabus demonstrates the syllabus week by week.

The study explores the correlation between the explicit teaching of the collocation and the lexical bundles and the improvement of the writing skill of First-year university students. The highlighted bold sections in every week demonstrated the explicit instruction of the collocations and lexical bundles.

The activities in the experimental group aim to explain to the participants the internal structure of the collocations and lexical bundles. The activities were designed to train the experimental group to look at learning the word differently. I trained them to understand how the collocation or lexical bundle is formed and the relationship between the bundles and the other word in the sentences. I also trained the experimental group to analyze the structure of the collocation and the

lexical bundles. The abovementioned was the first stage of noticing the collocation and the lexical bundles (Please refer to Appendix 4)

The noticing stage was followed by retrieving stage, during which the experimental group was given some reading articles and was asked to highlight any familiar collocations or lexical bundles. The retrieving stage is critical in learning because the participants need to understand that words do not exist in isolation. They are part of the text, and retrieving them and understanding their function in the text will have a positive role on the process of learning them (Please refer to Appendices 2 & 3)

The study focused mainly on writing skills, developing the sentence structure to writing well-developed pargraph, such as introductory paragraphs, body paragraphs, and conclusions. The participants were trained to write three genres: exemplification, argumentative, and cause and effect because they were part of the programme's learning outcomes. The final stage of the intervention study was the generating stage. The participants had to choose one of the three abovementioned genres to write their final essay. (Please refer to Appendix 7)

Before introducing the structure of the intervention scheme week by week in Table 2, it is worth mentioning that the rationale of the intervention study is designed to examine the correlation between the explicit teaching of the collocation and lexical bundles and the improvement of the writing skill of first-year university students. It is essential to define what I mean by explicit teaching. The study did not focus on in-class activities to mimic the collocation and lexical bundles. The end goal of the study is not for the students to memorize the list of collocations and lexical bundles by heart. On the contrary, the study aims to successfully train the students to use the taught collocations and lexical bundles in their essays.

To fulfill the abovementioned goal, I reviewed many studies that implemented collocations and lexical bundles in their teaching materials. Pellicer-Sánchez & Boers (2018) argued that the involvement with multiple instances with the target collocation and lexical bundles is important to make the collocations and lexical bundles memorable and understood. It is essential to have a pedagogical approach that gives the learners multiple encounters with the target language to assist them in learning the target language.

To ensure a successful learning experience of the proposed collocations and lexical bundles in the current study, I planned different in-class activities that introduced the learners to the target language before they used them in the essays at the end of the intervention study.

At the beginning of the study, I assigned simple in-class activities to ensure that the participants knew the complete list of the collocations and lexical bundles. Those activities were simply fill-in-the-blanks and multiple choices activities to assess the participants' knowledge of the collocations and lexical bundles (Please refer to Appendix 2- 5).

Later on, I introduced the participants to more challenging activities to ensure that they were capable of locating the taught collocations and lexical bundles in reading text. I assigned reading activities that would require the participants to underline the collocations and lexical bundles, enabling them to read how the collocations and lexical bundles create more comprehensible relationships between the sentences of the reading text. It is essential to combine two skills to train the students, and in this current study, the receptive skill of reading is crucial to expose the collocations and lexical bundles to them. The participants read the target collocations and lexical bundles in the reading passage as the first step before attempting to imitate them and regenerate them in their own texts at the end of the intervention study (Please refer to Appendix 6 for the reading practice in-class activities).

Lastly, the participants reach the generating process or the writing process. It is essential to realize that writing comes with many other skills: outlining, organizing, drafting, editing, and finally, writing the final draft. To have reliable results at the end of the intervention study, the study's participants encountered different types of in-class activities that introduced them to target collocations and lexical bundles. The in-class activities varied between choosing the correct collocations and lexical bundles to complete an essay. Another in-class activity required the participants to give peer review and edit each other essays to locate the mistake of the collocations and lexical bundles. (Please review Appendix 7 for the samples of the in-class writing practice).

Finally, the participants wrote their essays after encountering the target collocations and lexical bundles in various activities that challenged them and

directed their attention to the target language. The below schedule of the intervention scheme demonstrates the pedagogical approach towards the explicit teaching of the collocation and lexical bundles.

3.7 Participants

The study participants were first-year students at an American University in the Middle East who were enrolled in an English Language Bridge Prgramme. To improve their language proficiency, students who do not achieve the required grades on a TOEFL or IELTS enroll in the English Language Bridge Programme (ENGB). According to the Ministry of Higher Education in the UAE, students who do not score 550 out of 667 on the TOEFL test or 6.5 out of 9 on the IELTS cannot start their undergraduate programs. Students are thus placed in the English Language Bridge program (ENGB) to improve their writing skills for one semester or 15 weeks. The ENGB students spend four hours of classes weekly to improve their language level, thus joining the university and becoming eligible to take different courses in their chosen major. According to the university regulations where the study took place, all students should finish their English classes in the first academic year; otherwise, they will lose their offer or place in the university.

The participants in this study included 65 first-year students who scored below 6 in the IELTS or 550 in the TOEFL. Their scores place them in level B2 of the Common European Framework (CEF). The Common European Framework is a comprehensive tool describing different language levels (Heyworth, 2006). According to the CER, the B2 level learners are more independent English language users who can produce explicit language texts. The experimental group comprised 35 students from different countries in the Middle East with Arabic as their first language. The experimental group consists of 20 Females & 15 Males. They are all 19 and 20 years old.

The control group consists of 30 participants from the Middle East with Arabic as their first language. The control group consists of 15 Males and 15 Females. They are all between 19 and 20 years old.

The experimental group received training and instruction on collocation and lexical bundles, whereas the control group of 30 students did not. The two groups followed the same syllabus to achieve the same learning outcomes to fulfill the

course requirement; however, the experimental group had slightly different activities.

The participants learned about the intervention program through a flyer, which provided the time and location of the class. The participants voluntarily agreed to join the program to improve their writing skills. Once the participants volunteered, I gave them an ethical consent form and an information sheet explaining the programme's stages. I ensured that each population member had a fair chance of being involved in the sample. I also asked all the participants to suggest a pseudonym for themselves to write on the forms and informed them that they could withdraw from the study at any point during the 15 weeks.

Table 3
Outline of the Participants' Details

Participant ID	Age	Gender	Group
Participant 1	19 years old	Male	Experimental
Participant 2	19 years old	Female	Experimental
Participant 3	19 years old	Female	Experimental
Participant 4	20 years old	Male	Experimental
Participant 5	19 years old	Female	Experimental
Participant 6	20 years old	Male	Experimental
Participant 7	19 years old	Female	Experimental
Participant 8	20 years old	Female	Experimental
Participant 9	20 years old	Male	Experimental
Participant 10	19 years old	Male	Experimental
Participant 11	19 years old	Female	Experimental
Participant 12	20 years old	Female	Experimental
Participant 13	19 years old	Male	Experimental
Participant 14	19 years old	Male	Experimental
Participant 15	19 years old	Female	Experimental
Participant 16	19 years old	Male	Experimental
Participant 17	19 years old	Female	Experimental
Participant 18	19 years old	Female	Experimental
Participant 19	19 years old	Female	Experimental
Participant 20	20 years old	Male	Experimental
Participant 21	19 years old	Female	Experimental
Participant 22	20 years old	Female	Experimental
Participant 23	19 years old	Male	Experimental
Participant 24	19 years old	Male	Experimental
Participant 25	19 years old	Female	Experimental
Participant 26	20 years old	Male	Experimental
Participant 27	19 years old	Male	Experimental

Participant 28	19 years old	Female	Experimental
Participant 29	19 years old	Female	Experimental
Participant 30	20 years old	Female	Experimental
Participant 31	19 years old	Male	Experimental
Participant 32	19 years old	Female	Experimental
Participant 33	19 years old	Male	Experimental
Participant 34	20 years old	Female	Experimental
Participant 35	19 years old	Female	Experimental
Participant 36	20 years old	Female	Control
Participant 37	19 years old	Female	Control
Participant 38	20 years old	Male	Control
Participant 39	20 years old	Male	Control
Participant 40	19 years old	Female	Control
Participant 41	20 years old	Male	Control
Participant 42	19 years old	Female	Control
Participant 43	19 years old	Female	Control
Participant 44	20 years old	Male	Control
Participant 45	20 years old	Male	Control
Participant 46	19 years old	Female	Control
Participant 47	19 years old	Male	Control
Participant 48	19 years old	Male	Control
Participant 49	19 years old	Male	Control
Participant 50	20 years old	Female	Control
Participant 51	20 years old	Female	Control
Participant 52	19 years old	Female	Control
Participant 53	19 years old	Female	Control
Participant 54	20 years old	Male	Control
Participant 55	19 years old	Male	Control
Participant 56	19 years old	Female	Control
Participant 57	19 years old	Female	Control
Participant 58	19 years old	Female	Control
Participant 59	19 years old	Male	Control
Participant 60	19 years old	Male	Control
Participant 61	19 years old	Female	Control
Participant 62	20 years old	Male	Control
Participant 63	19 years old	Female	Control
Participant 64	19 years old	Male	Control
Participant 65	19 years old	Male	Control

3.8 Procedures for Data Collection

After receiving approval from the research ethics committee at the University of Exeter, I started the intervention study. The class was one hour long, four days a week. I began every class with a 15-minute revision of the collocations and lexical bundles explained in the previous session. I administered all the sessions during which the participants took their pre-tests and post-tests. I collected the tests and essays and kept them on my desktop. I stored the data on the University of Exeter One-Drive account with a secure password and data encryption during the study time. After the data analysis was finalized and approved, I immediately destroyed the data.

3.8.1 Data Collection Procedures

I designed all of the assessment tools in the intervention study to measure the correlation between teaching collocations and lexical bundles and the improvement in the writing skills of first-year college students.

I analyzed all of the pre-test and post-test results statistically using SPSS to examine the impact of teaching collocations and lexical bundles on the students' language development.

As a common practice in the English Department where I conducted the study, the final grades of all essays are determined by a panel of three English instructors who read the essays before grading them. After the panel finalized the grades, I highlighted all of the collocations and lexical bundles in the essays to examine whether the frequency of the occurrence of the formulas contributed to the overall grade. (Please refer to Appendix 14 for the complete data set of the study)

3.8.1.1 Data Preparation

The first part of data preparation involved comparing the pre-test and post-test scores between the experimental and control groups. I combined the *Experimental Group.xlsx* and *Control Group.xlsx* spreadsheets into a single spreadsheet, with the column header "Students" representing the student unique identifier (ID), "pre-test" representing the pre-test score, and "post-test" representing the post-test score. For consistency, I converted both pre- and post-

test scores to percentages by dividing by 30 and then multiplying the result by 100. To enable comparisons between the two groups, I generated a grouping variable "Group," with the value 1 representing the experimental group and 2 representing the control group. Finally, I calculated the outcome variable of "Change" (representing the change in score from pre- to post-test) as the post-test score minus the pre-test score for each participant.

Similarly, I combined the final scores from the essay (out of 100) into a spreadsheet, with pre-and post-test scores for the treatment and control groups for each participant. I took data for the control group from the spreadsheets *The control group pre intervention.xlsx* and *The control group post-intervention.xslx*, while I took data for the experimental group from the spreadsheets *The Experimental group pre intervention.xlsx* and *The Experimental group post intervention.xlsx***.** Again, a grouping variable for "Group" with the value of 1 represented the experimetal group and 2 represented the control group. Again, I calculated the outcome variable ("Change") as detailed above.

I only included participants with both pre- and post-test scores in the analyses aimed at comparing the outcomes across the two groups.

3.8.1.2 Data Analysis

I calculated the Mean, SD, Minimum, Maximum, and N (65 participants) for:

a. The pre- and post-test scores separately for the experimental and control groups.

b. The change in score between the pre- and post-test groups separately for the experimental and control groups using the results from the essays.

To check the distribution of the outcome measure (change in both tests and essay scores), boxplots were plotted to assess if the data are Normally distributed. The Shapiro-Wilk test of normality was used to test if the outcome was normally distributed. A p-value of less than 0.05 would indicate that the outcome variable was not normally distributed, in which case, non-parametric tests would be used to test for statistical significance. This will help determine what type of statistical tests are appropriate for the data.

3.8.1.3 Hypothesis Testing

I answered the following two questions through hypothesis testing as outlined below.

RQ1: To what extent do the explicit teaching of collocations and lexical bundles help students use them accurately?

To determine whether students could accurately use collocations and lexical bundles, I used two-sample t-tests. I conducted them separately for the teaching collocation and lexical bundles and essay scores from the overall writing grade.

- *Null hypothesis (H_0):* The true difference in mean changes in scores between the two groups is zero.
- *Alternate hypothesis (H_a):* The true difference in the mean changes in scores between the two groups is different from zero.

RQ2: Is there a correlation between the use of collocation and lexical bundles and the improvement of students' overall writing grades?

I created scatter plots and estimates of the correlation coefficients between teaching collocation and lexical bundles and the improvement of the students' overall writing grades separately for the intervention/treatment and control groups. For the correlations, I tested the following hypothesis for each group.

- *Null hypothesis (H_0):* Students' mastery of teaching collocation and lexical bundles and their overall writing grades are not significantly correlated.
- *Alternate hypothesis (H_a):* Students' mastery of teaching collocation and lexical bundles and their overall writing grades are significantly correlated.

Finally, I summed the frequencies in the use of Group A (Referential Expressions), Group B (Stance Expressions) and Group C (Discourse Organizing Functions) for all students to examine which expressions they used the most used on average. I then explored how the final grade in the overall writing test related to the commonly used expressions.

Q3. Is there a relationship between the increase in the number of lexical bundles used and the change in essay scores after intervention?

I calculated the total number of lexical bundles used in an essay for each student before and after the intervention for both the treatment and control groups.

I then calculated the change in the number of bundles by subtracting the number of bundles used before the intervention in each group from the number of bundles used after the intervention. I then tested correlation between the number of bundles used and the overall essay score based on the distribution of the data.

- *Null hypothesis (H_0):* The change in the number lexical bundles and overall writing grade are not significantly correlated.
- *Null hypothesis (H_a):* The change in the number lexical bundles and overall writing grade are significantly correlated.

3.9 Validity

I paid a great deal of attention to controlling aspects that would affect the validity of my research. The external validity in quantitative data is very much concerned with the generalizability of the study, or how the sample study can be generalized to a population (Cohen et al., 2007). I also needed to consider the lack of representation of the available target population and ensure that my population represented the population for which I sought to generalize the results. I paid attention to instrumental viability or unreliability when planning or choosing the collocation and lexical bundles. Internal validity is concerned with whether the question of the proposed study makes a difference. There were many threats to the internal validity of the study, including selection bias. As such, I ensured that my population was randomly selected (Creswell, 2008).

It is crucial that any measuring instrument /test measures what it is designed to measure for the results of the study to be valid (Surucu & Maslakci , 2020). The validity of any experiment depends on how well the measuring instruments perform is function (Anastasi & Urbina, 1997). Validity is also determined by the interpretation and the results of the data taken from the tests administered in the study.

It is imperative to consider construct validity, especially in experimental studies during which an assessment is required to report data at the end of the study. In the current study, the assessment involved a pre-test and post-test that determined students' knowledge of collocations and lexical bundles. Before developing the studies tests, it was very essential to review all the previous attempts to develop tests of collocations and lexical bundles. The literature review

chapter contains a complete account of the tests in section 2.6. Gyllstad (2007)gave a full review of how Formulaic language testing had developed over the years and suggested that psycholinguistic techniques might help provide more valid tests. Surprisingly, Schmitt (2022) argued that regardless of the long list of tests that were developed to test some aspects of the Formulaic language, they still do confirm the learner's collocational knowledge.

Despite the above, the current study needed to develop pre-test and post-test to measure the knowledge of the collocations and lexical bundles. To fulfill the task, I followed Gyllstad & Schmitt (2019) to develop tests that ensure some validity. I had to build a construct through my explicit teaching of the target collocations and lexical bundles, besides providing a fundamental purpose for the tests. I explained the purpose of every fill-in-the-blanks test and the multiple-choice test. I also provided clear instructions on the formatting of the tests and why the multiple choice/fill-in-the-blank test is out of 30 and the writing is out of 100. It is essential to provide transparency and explain the grading system clearly to ensure the validity of the tests.

One of the essential aspects to consider when having tests as the measuring instruments of the research is content validity using statistical methods. In this study, I used pre-test and post-test of multiple choice and fill- in -the blank tests, besides the results of the writing tests, which were aggregated in a table of pre- and post-writing test scores. I used factor analysis which Charles Spearman adopted in early 1900. The Factor analysis uses mathematical procedures to explore the patterns in the set of variables in the study. The factor analysis summarizes the data to be easily understood and explains the relationship between the variables of the tests (Surucu & Maslakci, 2020). Please refer to Appendix (11) for a possible influence of outliers on the correlations.

Additionally, reliability should be considered when reporting the results of any experimental studies. Reliability means the stability and consistency of the tests that are used in the study. Reliability means the ability to redo the same tests and gets the same results. Indeed, this is not possible due to the change in the circumstances around the test, mainly since the current study measures the participants' language ability. It is not possible to get the same results if the test is repeated because the study population will be different. To ensure the reliability of

the test results, I considered specific measures, especially in reporting the writing assessment test results. All the results of the essays had to be determined via a rubric (please see Appendix 8). The writing pre-test and post-test grades were determined by a committee of three English language instructors who are full-time faculty at the university where the study took place. The committee grades all the essays anonymously, and they must decide the writing grade out of 100.

3.10 Ethical Consent

I sent an informed ethical consent form and an information sheet explaining the purpose of the study to all the participants prior to beginning the study. I ensured the confidentiality of the participants' identities throughout the research by giving each student a pseudonym to anonymize them. The essays and tests only contained a number, not a name. I also ensured non-traceability throughout the study, which I extended to the aggregating of the data. I did not share the data with third parties except with my supervisors. I treated the information shared with the strictest confidence and did not divulge it without the participants' permission or for reasons beyond the purposes of this study.

Before I started my study, I created my profile on the workribe platform to start an ethical application. I uploaded all the required documents from the Participants' information sheet and consent form to get the FHASS social science and international studies ethics committee approval. After the committee approved my ethics application, I started my study. Please refer to Appendices 12 & 13 for the ethical approval, participants' information sheet, and consent form.

CHAPTER 4: Findings

In this chapter, I investigate the influence of using collocations lexical bundles on test and essay scores in a group of first-year university students in the English Language Bridge Program. Students were classified into two groups (experimental and control) and were assessed using a test and an essay before and after the intervention. The scores in the test and essays were then compared between the experimental and control groups. Comparisons were made for:

1) Scores in the experimental and control groups. These were called pre-test. scores.

2) Scores in the experimental and control groups after the experimental group had been trained to use the lexical bundles. These were called post-test scores.

To investigate the effect of the intervention (use of lexical bundles) on student performance, the change in score for both the test and essay was calculated as the difference between the post and pre-intervention scores for each student. A change greater than zero would imply that the student's performance on the test or the essay

improved after the intervention.

I calculated the Mean, SD, Minimum, Maximum and N (65 participants) for:

a. The pre and post-test scores separately for the experimental and control groups

b. Change in score between the pre and post-test scores separately for the experimental and control groups using the results from the essays

To check the distribution on the outcome measure (change in both test and essay scores), boxplots were plotted to assess the distribution of the data and also to graphically compared the test and essay scores before and after the intervention in the two groups. This determined what type of statistical tests were appropriate for the data.

4.1 Descriptive Analysis

This study included a total of 65 participants, with 35 participants assigned to the experimental group and 30 to the control group. The mean (SD) of the pre-

test score was 44.7 (7.3) in the experimental group, while students in the control group had a slightly lower mean grade 40.1 (8.0) (see Table 3 and Figure 1). The post-test scores in the experimental group were markedly higher, 66.7 (12.7), compared to the control group, 41.1 (10.4), which achieved the same scores as in the pre-test, on average.

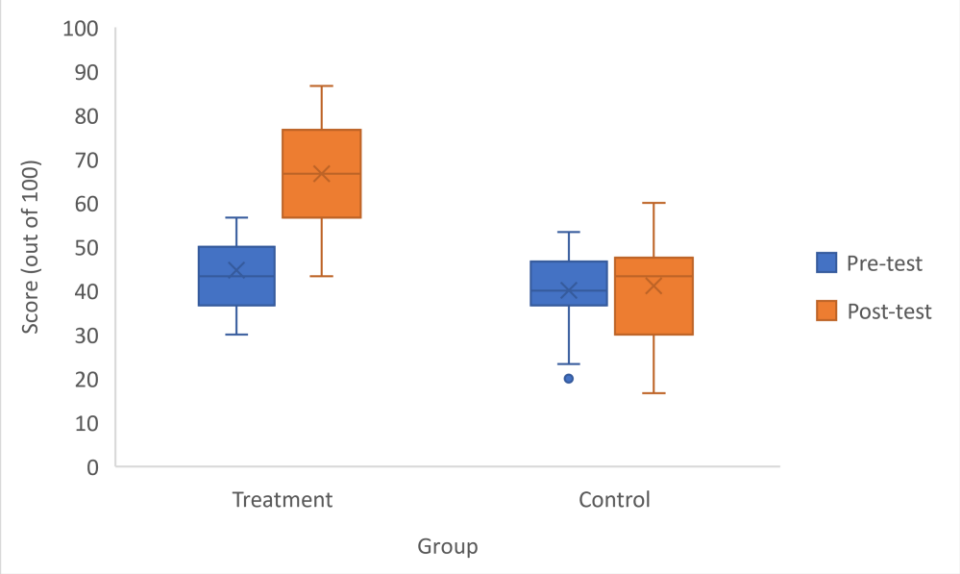
Table 4

Descriptive Summary of Control and Experimental Group Test and Essay Scores by Group

	Group	N	Mean (SD)	Median (IQR)	Minimum	Maximum
Summary of Test Scores by Group						
Pre-test	Experimental	35	44.7 (7.3)	43 (37, 50)	30	57
	Control	30	40.1 (8.0)	40 (37, 47)	20	53
Post-test	Experimental	35	66.6 (12.7)	67 (57, 77)	43	87
	Control	30	41.1 (10.4)	43 (30, 47)	17	60
Summary of Essay Scores by Group						
Pre-test	Experimental	35	76.1 (3.0)	77 (73, 79)	70	80
	Control	30	76.6 (4.0)	77 (75, 78)	66	83
Post-test	Experimental	35	80.7 (4.6)	82 (76, 85)	72	87
	Control	30	77.9 (4.0)	79 (77, 80)	66	83

Figure 1

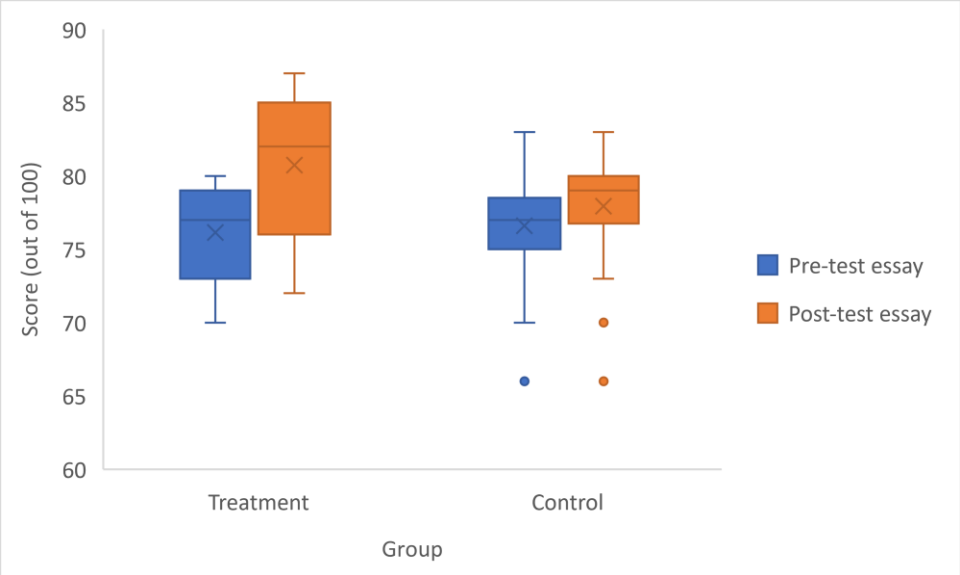
Boxplots of Pre-Test and Post-Test Scores by Group



There was a greater improvement in the essay scores among those assigned to the intervention (see Table 3 and Figure 2). I used Boxplots to check the distribution of the scores and to graphically show if there were any differences in the distribution ion of the pre-test and post -test scores.

Figure 2

Boxplots of Pre- and Post-Intervention Essay Scores by Group



4.2 RQ1: To What Extent Does Teaching Collocations and Lexical Bundles Help Students Use Them Accurately?

4.2.1 Comparing Pre- and Post-Test Scores by Group

On average, the mean post-test scores in the experimental group increased by 22%, mean (SD) 22.0 %(7.9), while in the control group, the mean (SD) change in the test score was 1.0 (9.1). To check whether there were differences in the change in scores between the two groups, I first calculated the Shapiro-Wilk p-value to test whether the changes in scores were normally distributed in both groups (see Table 4 and Figure 3).

Table 5

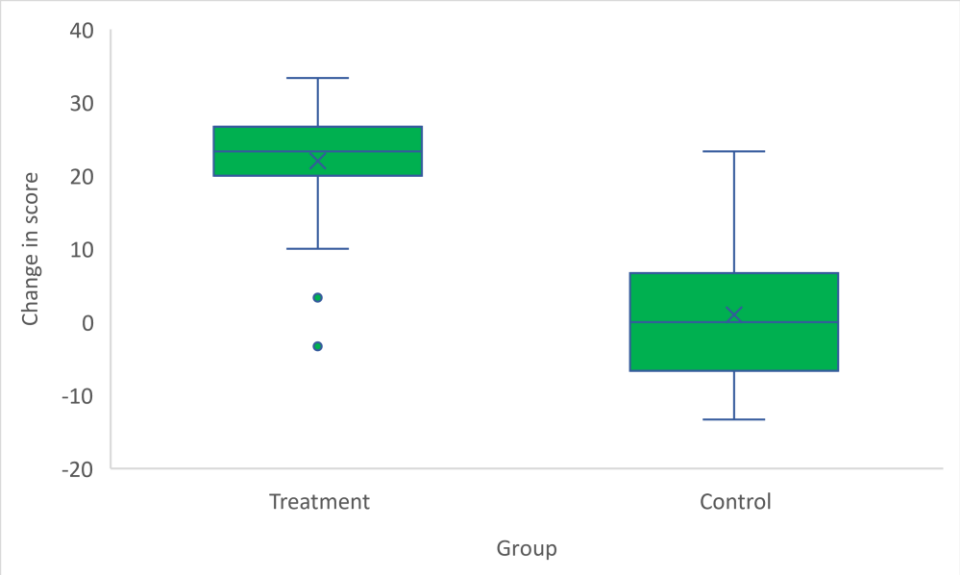
Descriptive Summary of the Change in Test Scores by Group

Group	N	Mean (SD)	Median (IQR)	Minimum	Maximum	Shapiro-Wilk test for normality p-value
Experimental	35	22.0 (7.9)	23 (20, 27)	-3	33	0.002
Control	30	1.0 (9.1)	0 (-7, 7)	-13	23	0.118

The distribution of the change in scores in the experimental group did not satisfy the normality assumption based on the Shapiro- Wilk test for normality , p=0.002. I thus used the non-parametric Wilcoxon rank-sum (Mann-Whitney U test) to test whether there were no differences in the change in scores between the two groups. The p-value from the Mann-Whitney U test was < 0.05. I thus rejected the null hypothesis that difference in the mean change in scores in the two groups was 0, in favor of the alternative that the difference in the mean change in scores between the two groups was different from 0 following the intervention. As shown in Table 4 and Figure 3, the change in scores in the experimental group was bigger than that observed in the control group.

Figure 3

Boxplots of the Change in Score by Group



4.2.2 Comparing Pre- and Post-Intervention Essay scores by Group

The mean (SD) change in essay scores were 4.6 (3.0) and 1.3 (3.4) in the Experimental and control groups, respectively (see Table 5 and Figure 4).

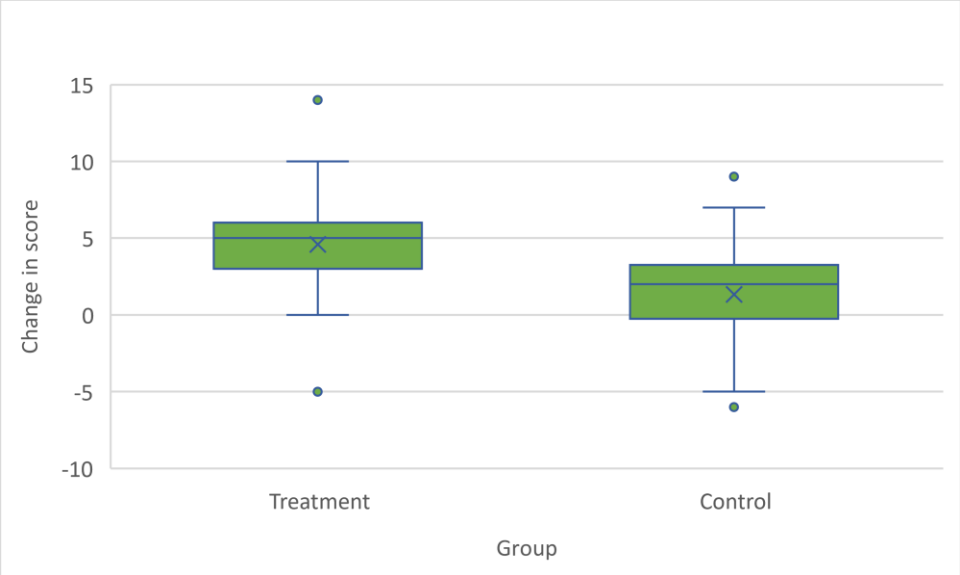
Table 6

Descriptive Summary of the Change in Essay Scores by Group

Group	N	Mean (SD)	Median (IQR)	Minimum	Maximum	Shapiro-Wilk test for normality p-value
Experimental	35	4.6 (3.0)	5 (3, 6)	-5	14	0.001
Control	30	1.3 (3.4)	2 (0, 3)	-6	9	0.748

Figure 4

Boxplots of the Change in Essay Scores by Group



Again, the distribution of the change in essay scores in the experimental group did not satisfy the normality assumption, p=0.001. I used the non-parametric Wilcoxon rank-sum (Mann-Whitney U test) again to test whether there were no differences in the change in the essay scores between the two groups. The p-value from the test was < 0.05. I thus reject the null hypothesis that the difference in mean change in the essay scores in the two groups was 0, in favor of the alternative that the difference in the mean change in essay scores between the two groups was different from 0 following the intervention. As shown in Table 5 and Figure 4, the mean change in essay scores in the experimental group (4.6%) was bigger than that observed in the control group (1.3%). In the Experimental group, all but two essay scores from two individuals were between 0 and 10 in the experimental group.

4.3 RQ2: Is There a Correlation Between Teaching Collocation and Lexical Bundles and the Improvement of Students' Overall Writing Grades?

4.3.1 Correlation Between the Change in Test and Essay Scores

I plotted scatterplots with accompanying estimated correlation coefficients between the changes in essay scores against the changes in test scores separately for each group to ascertain if there was a relationship between a

student's change in test score and their overall grade in the essay exam (see Figures 5 and 6).

The estimated Spearman correlation coefficient between the change in the test and essay scores among the controls was -0.06 (95% CI: -0.41, 0.29), indicating a weak negative relationship between the scores in the test and in the essay. This estimated correlation was not significantly different from 0 (p-value = 0.718). I thus rejected the null hypothesis that there was a significant correlation between the test and essay scores among the controls.

For those assigned to the experimental the scatterplot showed that there was a positive relationship between the test and essay scores (see Figure 6). The estimated Spearman correlation coefficient between the change in the test and essay scores in the experimental group was 0.51 (95% CI: 0.23, 0.80), and this was significant (p-value < 0.001). This implies there was a significant positive relationship between the test and essay scores among those assigned to the intervention.

There was a negative relationship between the changes in test and essay scores among the controls and a positive relationship between the two scores in the experimental group (see Figures 5 and 6).

Figure 5

Scatterplot of Changes in Scores (Essay) by Change in Score in the Control Group

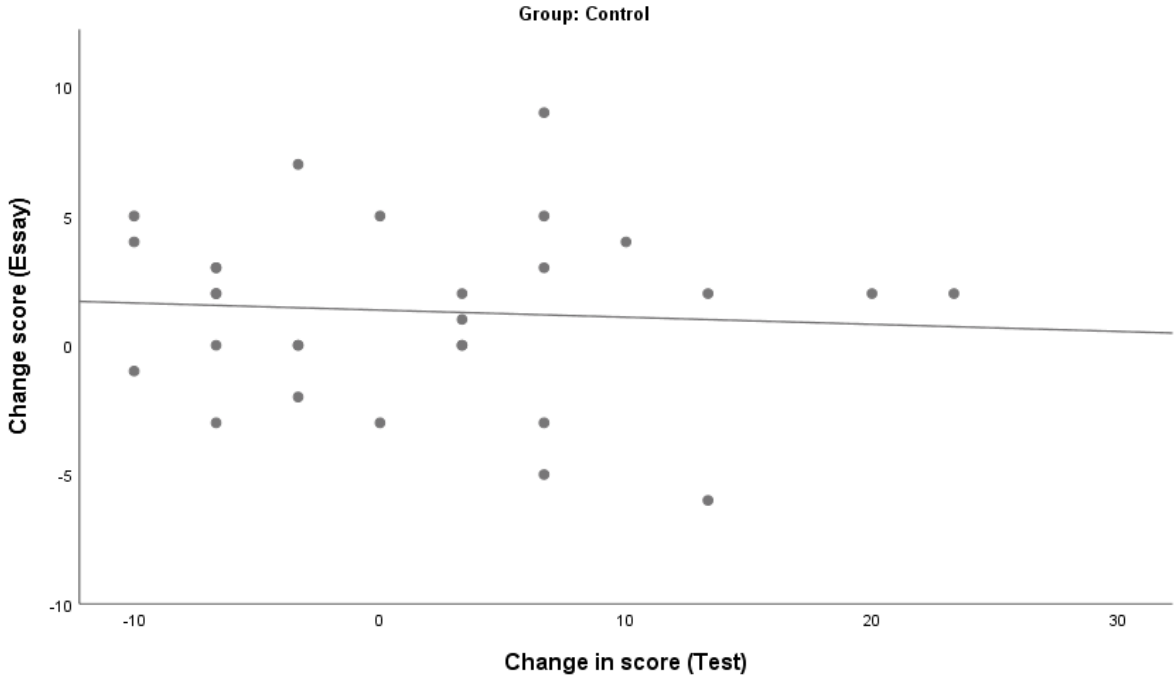
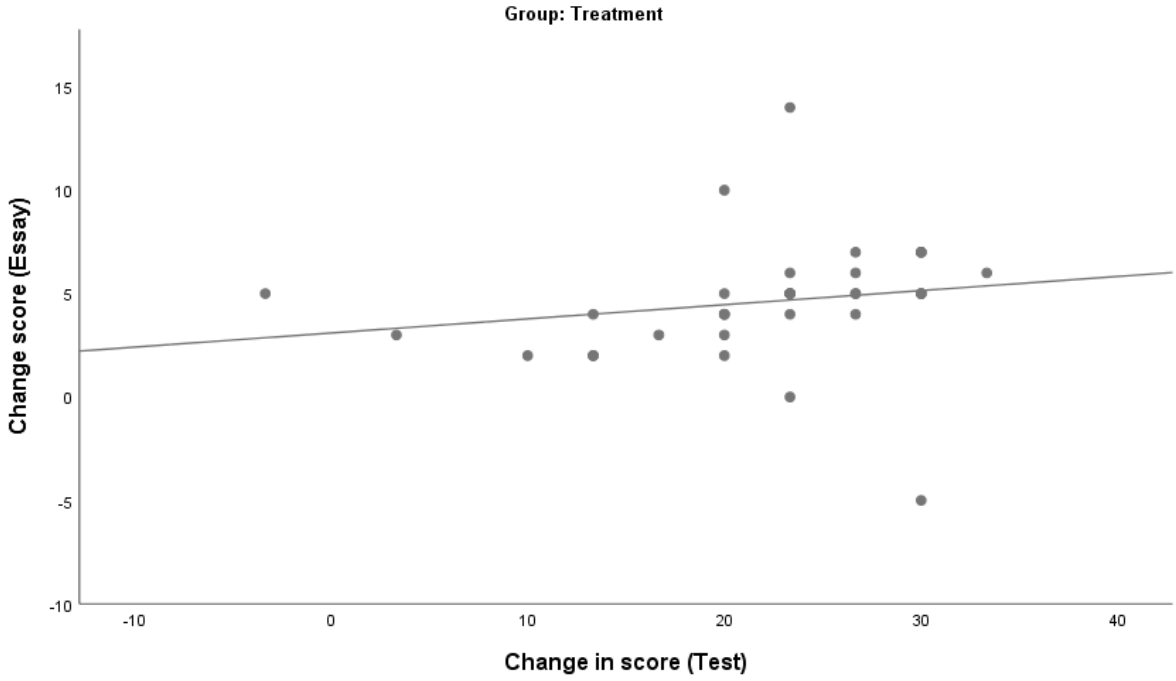


Figure 6

Scatterplot of Changes in Scores (Essay) by Change in Score in the Experimental Group



4.3.2 Changes in the Use of Lexical Bundles Following Intervention

I tallied the lexical bundles used in writing separately for the two groups to highlight the most used bundles before and after the intervention. In the experimental group, *doesn't have*, *as a whole*, *at this stage*, *in a similar manner*, and *nevertheless* were in the top 20 most used bundles prior to the intervention, with almost all of the students using them at least once in their writing (see Table 6). After the intervention, *due to the fact*, *on the other hand*, *the degree to which*, *this means that*, and *to distinguish between* were the most frequently used bundles in the experimental group, with participants using them at least once in their write-ups, on average.

For the control group, *an attempt to*, *doesn't have*, *at this stage*, *it follows*, and *is affected* were some of the most commonly used bundles prior to the intervention, with each participant using them once on average in their write-up (see Table 7). Post-intervention, *as a result* and *doesn't have* were the most commonly used bundles among the controls, with students using them around two times on average. *Nevertheless*, *this does not*, and *in a similar manner* were the other most commonly used bundles in the control group.

Table 7

Use of Lexical Bundles in the Experimental Group Showing 20 of the Most Used Bundles

Pre-intervention		Post-intervention	
Bundle	Times used by all 35 students	Bundle	Times used by all 35 students
doesn't have	34	due to the fact	63
in a number of	34	on the other hand	60
as a whole	34	degree to which	55
at this stage	33	this means that	54
in a similar manner	32	to distinguish between	54
nevertheless	30	it has been	53
on the other hand	23	are able to	53
if this is	22	to determine whether	53
there are several	22	it is necessary	53
this means that	22	with regard to	52
can be expressed	22	there was no significance difference	50
be argued that	22	it follows	49
most likely to	22	at the end of this point	48
it follows	22	which is more	48
is affected	21	it is worth	48
An attempt to	20	doesn't have	47
some people might argue	17	it can be	47
as a consequence	15	it is obvious that	47
are/was based on	14	the effects of	47
this does not	14	is affected	46

Table 8

Use of Lexical Bundles in the Control Group Showing 20 Most Common Bundles

Pre-intervention		Post-intervention	
Bundle	Times used by all 30 students	Bundle	Times used by all 30 students
An attempt to	30	as a result	68
doesn't have	30	doesn't have	61
at this stage	30	nevertheless	52
if this is	30	this does not	50
		in a similar	
there are several	30	manner	50
in a number of	30	in contrast	32
this means that	30	on the other hand	32
can be			
expressed	30	at this stage	30
be argued that	30	there are several	30
most likely to	30	as a whole	30
on the other			
hand	30	An attempt to	29
as a whole	30	if this is	29
it follows	30	can be expressed	29
is affected	30	due to the fact	29
nevertheless	29	be argued that	28
in a similar			
manner	29	most likely to	28
are able to	2	it follows	28
		is affected	28
		in a number of	27
		this means that	27

4.4 RQ3: Is There a Relationship Between the Change in the Number of Lexical Bundles Used and the Change in Essay Scores After Intervention?

For this part, I tallied the number of times each participant used each of the lexical bundles to determine the total number of times each student used the bundles in their write-up before and after intervention for both groups. The number of bundles used in writing in the experimental group changed from 19.3 (4.5) to 89.0 (46.6) following the intervention, while for the controls, the mean number of bundles slightly increased from 16.0 (0.5) to 24.5 (5.2) (see Table 8 and Figure 7).

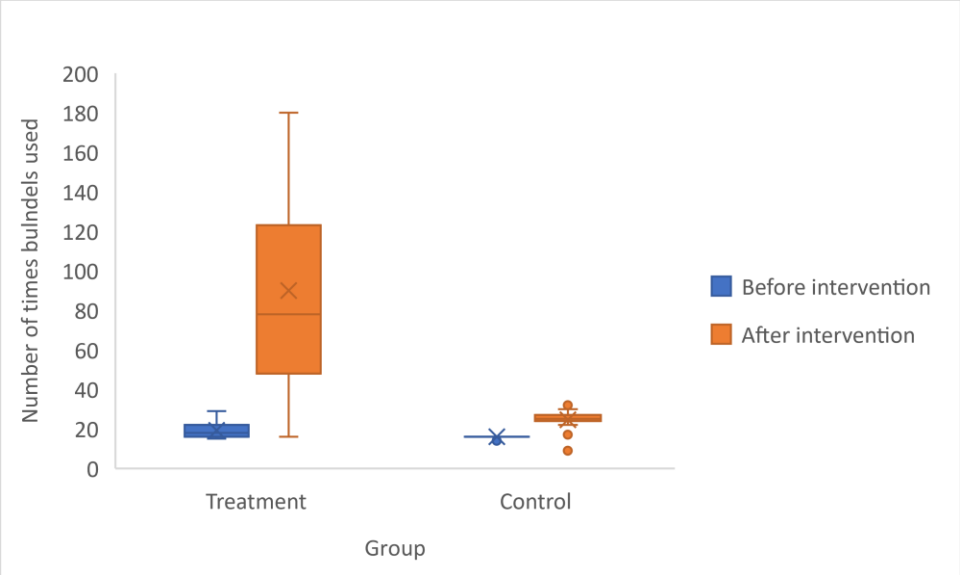
Table 9

Descriptive Summary of Use of Lexical Bundles in Essays Pre- and Post-Intervention by Group

Group	N	Mean (SD)	Median (IQR)	Minimum	Maximum
Before Intervention					
Experimental	35	19.3 (4.5)	18 (16, 22)	15	29
Control	30	16.0 (0.5)	16 (16, 16)	14	18
After Intervention					
Experimental	35	89.9 (46.6)	78 (48, 123)	16	180
Control	30	24.5 (5.2)	25 (24, 27)	9	32

Figure 7

Boxplot of Number of Lexical Bundles Used Before and After Intervention by Group



The number of bundles used in writing increased by 70 on average in the experimental group, compared to an average increase of 8.5 bundles in the control group (see Table 9). I then estimated the Spearman correlation coefficient between the change in the number of bundles used and the change in the essay scores because the change in bundles in the control group did not satisfy the normality assumption.

Table 10

Descriptive Summary of the Change in Number of Lexical Bundles Used in Essays

Group	N	Mean (SD)	Median (IQR)	Minimum	Maximum	
Experimental	35	70.6 (46.6)	59 (32, 107)	-2	156	0.109
Control	30	8.5 (5.1)	9 (8, 11)	-7	16	<0.001

The correlation between the change in number of bundles and change in essay score after intervention in the control group was 0.19 with 95% CI -0.18 to 0.56, which was not significant (p = 0.308) (see Figure 8). On the other hand, there was a strong positive correlation between the change in the number of bundles and the change in essay scores after intervention in the experimental group in the experimental group, 0.72 (95% CI: 0.55 to 0.90, p-value < 0.001) (see Figure 9).

Figure 8

Scatterplot of Changes in Scores (Essay) by Change in Number of Bundles Used in the Control Group

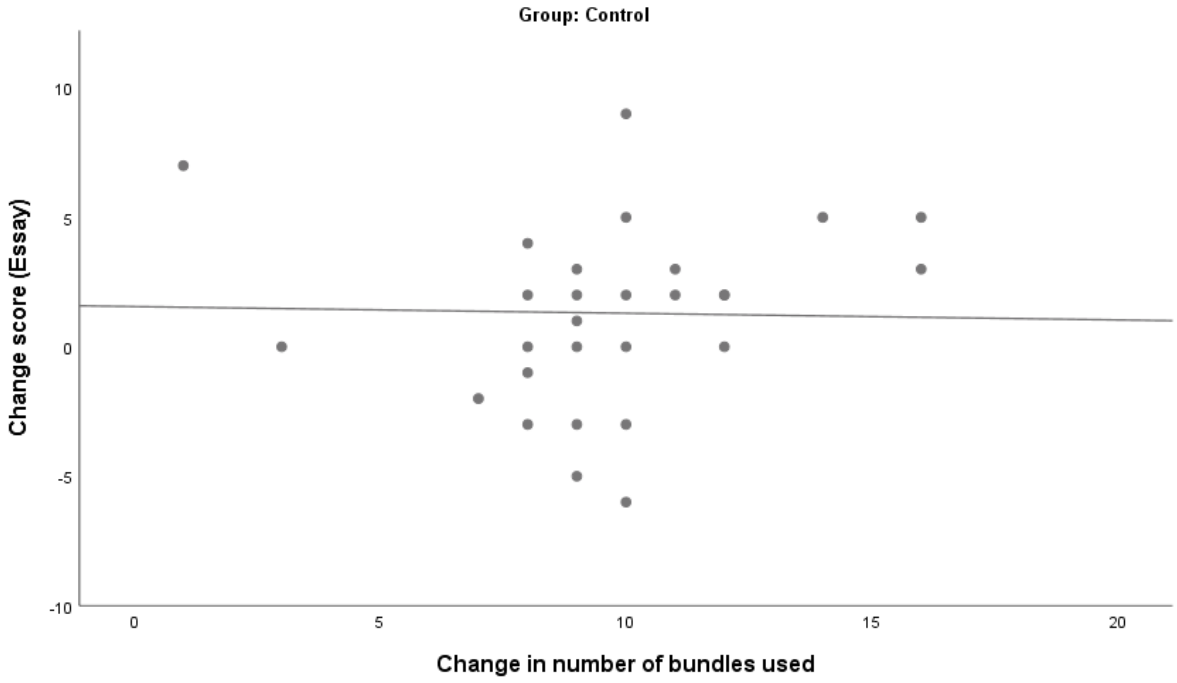
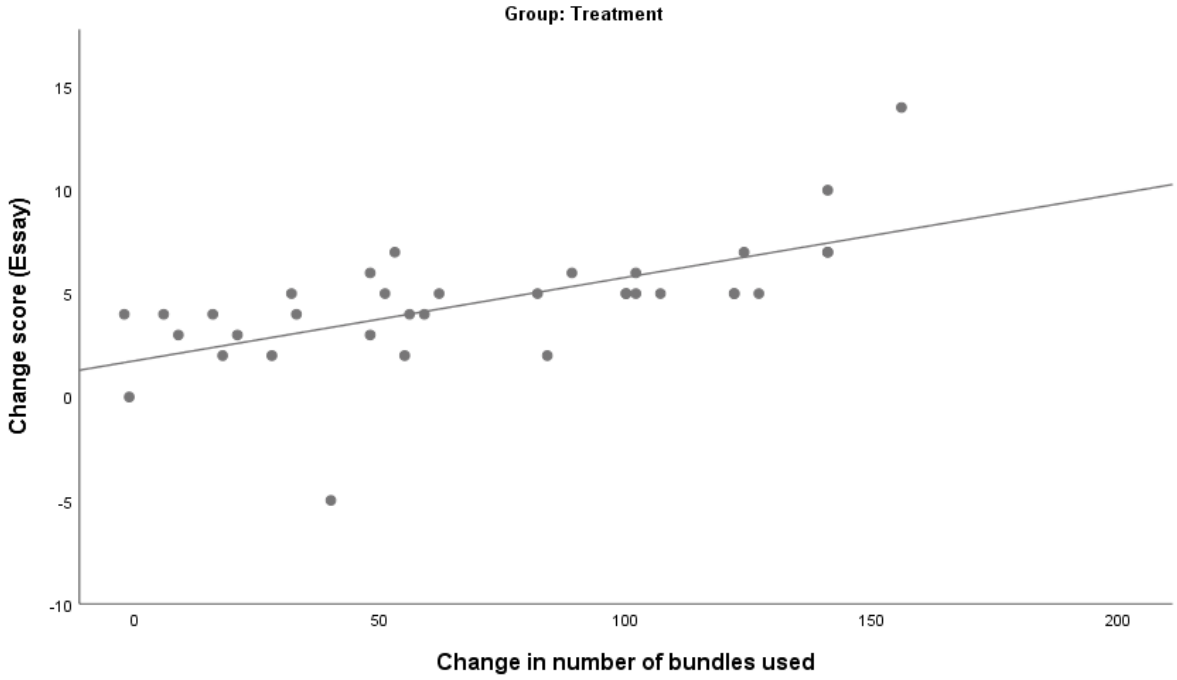


Figure 9

Scatterplot of Changes in Scores (Essay) by Change in Number of Bundles Used in the experimental Group



4.5 Summary

- Participants in the experimental group achieved a bigger change in their test and essay scores after the intervention compared to the controls.
- There was no linear relationship between the test and essay scores in the control group, while I found a significant correlation between the test and essay scores in the experimental groups.
- I identified some commonly used lexical bundles in both groups, the use of which generally increased among all students. I have listed them in Table 6 & 7.
- The number of lexical bundles used in writing was positively correlated with an improved grade in the essay scores in the experimental group.
- Students who used more bundles tended to score higher marks than those who did not.

CHAPTER 5: Discussion

This chapter discusses the findings of the present study and explains further how the current study differs from previous studies that have investigated the impact of using collocations and lexical bundles on the improvement of the language proficiency, especially writing skills. This chapter theorizes the findings and relates the results to existing research.

5.1 Discussion of the Research Findings

The study examined the correlation between the teaching collocations and lexical bundles and the improvement of first-year university students' writing grades. It also examined to what extent the explicit teaching of collocations and lexical bundles helps students use them accurately. The study investigated the relationship between the increase in the number of lexical bundles and the difference in overall writing grades. The third research question aimed not only to question the effectiveness of collocations and lexical bundles but to highlight the groups of collocations and lexical bundles that the students used most commonly to determine whether those bundles should be recommended to enhance the writing skills of first-year university students.

The study did not question if collocations and lexical bundles play a role in learning a foreign language; this question has been discussed in several studies over the last three decades. Teaching English, especially the EFL field, has benefitted from the introduction of formulaic language, in general, with collocations and lexical bundles, in particular.

The study's central question was not to argue the vital role of collocations and lexical bundles in learning English. Conversely, the study investigated what type of collocations and lexical bundles would be more practical and effective in learning English as a foreign language, particularly in developing students' writing skills. As such, the study focused on academic writing, which is one of the essential EFL productive skills.

The study was designed in stages that built on each other. It first asked, if we explicitly teach collocations and lexical bundles in EFL classes, will the students be able to use them accurately in tests, like fill-in-the-blanks and multiple-choice

tests? Second, does students' knowledge of individual collocations and lexical bundles eventually help them write better essays? Would this knowledge eventually improve their overall writing grades? Lastly, if the students showed progress in their writing skills, what are the types of collocations and lexical bundles they used in their writing?

When discussing collocations and lexical bundles, we cannot refer to them as a single entity or maintain a holistic view of them. We need to understand that, under the umbrella of collocations and lexical bundles, there are many types of each, and the question is what types of collocations or lexical bundles help students most and why these types are more beneficial when it comes to writing.

Several studies have investigated the relationship between using collocations and lexical bundles and improving writing skills. Staples et al. (2013) studied the use of collocations and lexical bundles in TOEFL writing tasks and concluded that the lower the students' level, the more they used lexical bundles and collocations. Staples et al. (2013) also noted that the students tended to recycle the collocations and lexical bundles they found in the question prompts. Appel and Wood (2016) confirmed these findings, noting an increase in the use of collocations and lexical bundles in the low-scoring writing of students in the Canadian Academic English Language Assessment (CAEL). Appel and Wood's (2016) model found that the low-scoring students made use of stance bundles and discourse organizing bundles, whereas high-scoring students used more referential bundles

By comparing my findings with those of Staples et al. (2013) and Appel and Wood (2016), It is soon to confirm a direct relationship between using collocations and lexical bundles and improving students' writing skills. However, the present study revealed that lower-scoring students tend to use more two or three -word stance bundles and discourse bundles, such as *even though*; as *a result*, *it is necessary*, and tend to repeat them many times in every paragraph. For instance, if the student is writing three body pargraphs, they use them in every body paragraph and sometimes in identical sequences, which makes the writing monotonous in style and robotic. On the other hand, high-scoring students use more referential bundles, which can be molded into the sentences rather than serving as fixed lexical bundles. As such, these require a certain level of

proficiency for students to use them in sentences. Table 6 in the “Findings” chapter lists the most common lexical bundles of post-intervention writing. The high-scoring students in the Experimental group used many referential bundles. In contrast, the control group used stance and discourse bundles that they used without any change in form, as mentioned in Table 7 in the “Findings“ chapter.

Durrant (2018) argued that scholars’ understanding of the relationship between using some formulas and improving writing quality still has a long way to go. However, there is an apparent increase in the demand to include more collocations and lexical bundles in the English language, particularly in EFL classes. Nevertheless, we should not encourage the production and circulation of ready-made lists of collocations and lexical bundles because this approach to teaching English does not help students to improve their language skills. On the contrary, it trains students to depend on ready-made lists and to use them in their writing without trying to improvise or diversify the lexical bundles. Pellicer - Sanchez & Boers (2018) emphasized the significance of adopting a pedagogical approach that provides learners with multiple encounters with the collocations and lexical bundles to assist them in learning the target collocations and lexical bundles.

For the above recurring need to investigate an appropriate pedagogy that could work well while teaching the collocation and lexical bundles as per Pellicer - Sanchez & Boers (2018) recommendations, it was essential to incorporate the taught collocations and lexical bundles into communicative classes as demonstrated in Table 2. The intervention study adopted a communicative teaching pedagogy where every taught language formula, collocation, or lexical bundle was explicitly taught through interactive communicative lessons, The current study implemented the explicit instruction of collocation and lexical bundles. The current study was influenced by Jones and Haywood’s (2004) model because their model sheds light on the relationship between the explicit instruction of formulaic language and the development of language proficiency. Jones and Haywood (2004) did not simply introduce formulas to the participants; on the contrary, they taught the participants some strategies to learn the formulae to transfer the formulas into acquired knowledge. This model helped raise students’ awareness of the sequences so they could learn the formulas through explicit

instruction. In my current communicative study, I also incorporated Nation's (2001) process, which explained that learning new words depends on three psychological processes: noticing, retrieving, and generating. Jones and Haywood (2004) paid attention to the little details of formula learning. What made their model good was their understanding that to learn a formula, students need to notice it as they read, retrieve it through tasks, and generate it through writing activities. These steps are all critical for the formula to be considered acquired knowledge.

The communicative language teaching pedagogy that was used to teach the collocations and lexical bundles in the current study resulted in students scoring higher writing grades in the experimental group. The scores of the high-achieving students demonstrated that using the attitudinal type of collocations and lexical bundles contributed to developing better sentence structure in the experimental group. Through the use of attitudinal collocations and lexical bundles, the high-achieving students express their evaluations and attitudes, besides the explicit teaching of attitudinal lexical bundles helped the high-achieving students to build stronger arguments and engage the reader by using various types of lexical bundles; such as it *should be noted that*, *as can be seen*, *it is possible to*, *that fact that the*, and *are more likely to* (Dontcheva-Navratilova,2012).

Collocations and lexical bundles should be regarded as a language learning tool, not an end. They should be utilized to inspire learning, understand different forms of sentence structure, and not create passive learners who use ready-made bundles in their writing without making any alterations. Learning collocations and lexical bundles should promote acquired knowledge rather than temporary learning experiences. For all the above-mentioned reasons, teachers who aspire to incorporate collocations and lexical bundles should do so via communicative pedagogy that introduces the taught collocations and lexical bundles via interactive communitive language materials.

On the explicit instruction of formulaic language, especially collocation, Durrant and Schmitt (2010) carried out an experiment on non-native ESL students, during which they introduced the participants to some collocations in reading comprehension. They instructed the participants to read some low-frequency and adjective-noun collocations out loud. They then tested the participants to examine whether they could recall the collocations and found that the students remembered

the collocations they had read. Although the Durrant and Schmitt (2010) study took place in a classroom with timed activities, it is still considered a successful model of how the explicit instruction of collocational knowledge helped students learn new collocations (Sonbul & Schmitt, 2013).

The current study confirmed the benefits of the explicit instruction of collocations and lexical bundles to non-native speakers (Schmitt, 2008). Although many studies have examined the impact of explicit instructions on native and non-native language learners, we cannot treat or examine both groups with the same tools because the language exposure of non-native learners cannot compare to that of native speakers. Hoey (1991-2005) introduced collocational priming and argued that collocations are acquired incrementally over time when students are exposed to them through input, and that such exposure might help create new collocations or modify old ones. Hoey (2005) also noted that there is little difference between the collocational priming of natives and non-natives. Hoey(2005) argued that non-natives might encounter collocational knowledge from their surroundings and benefit from indirect input, while non-natives, given their circumstances, might not have the same exposure as native speakers but benefit more from explicit instruction on collocational knowledge, first through clear instructions on collocations and then through activities and tasks (Sonbul & Schmitt, 2013).

The present study investigated how the use of collocations and lexical bundles contributed to improving the writing of first-year EFL university students. To do so, I manually wrote down the collocations and lexical bundles the high-achievement participants used in the experimental group. By reviewing the essays of the experimental group, I noticed that the students who achieved better grades were those who used more of the lexical bundles that included noun phrases with other modifier fragments and verb phrase + *that* clause fragments (see Table 6).

The study results confirm that the experimental group students benefited from the taught collocations and lexical bundles. By reviewing the data of the high-achieving students in the experimental group, it is evident that there is a relationship between the explicit teaching of the collocations and lexical bundles and the improvement of the overall writing grade. Improving the writing grades should be considered a step in the right direction, which should be followed with

more research. We can use the current results to propose that raising awareness of using collocations and lexical bundles, as Boers and Lindstromberg (2012) noted, helps high-achieving students write more complex sentences rather than resort to short sentences. The focus should now be on how lexical bundles help EFL learners improve their sentence structure and write more complex sentences to express their ideas better. I will discuss in detail some recommendations for future research in section 5.3.

As a result, this study showed that not only the use of collocations and lexical bundles helped the students in the experimental group achieve better grades than the control group. On the contrary, what mattered was the type of lexical bundles and how the students in the experimental group used them to fulfill their tasks and write their sentences. The high-achieving participants in the experimental group achieved better grades because of the way they used lexical bundles; they did not just use them in their simplest form, but they molded and changed them to suit their sentence structure and expressive needs. After the committee of English teachers determined the overall grades of the experimental group, I analyzed the type of lexical bundles that the high-achieving students in the experimental group used. I concluded that the high-achieving students scored better in writing grades because they used the lexical bundles to write more complex sentences. They used the bundles in Table 6 to write more complex sentences with multiple clauses that demonstrated better command of the English language. The complexity of the sentence structure and the variety of clauses in the sentences of the high-achieving student of the experimental group allowed them to express their ideas better in a more coherent and advanced structure, which improved the writing quality of the essays and eventually resulted in better overall grades.

There has been increased research on the impacts of teaching lexical bundles on improving writing skills. Most studies have agreed on the importance of introducing more lexical bundles in EFL classes, finding that they positively impact students' writing skills. A study in Indonesia investigated the type of lexical bundles students used in academic writing classes. It concluded that the participants were aware of the importance of using lexical bundles to improve their writing (Sugiarti et al., 2018). Sugiarti et al. (2018) also found that the participants tended to use more

research-oriented lexical bundles in the result and discussion sections of their academic writing.

There is a clear increase in demand for studies on the impact of lexical bundles, but language researchers should maintain a realistic view when discussing the impact of lexical bundles, as they are not a magic formula that can improve language acquisition. Despite their clear importance to writing, studies have noted that acquiring lexical bundles requires a lot of time and exposure (Cortes, 2004, 2006). They are acquired incrementally and over an extended period, just like single words (Nation, 2001; Schmitt, 2000).

The present study investigated the correlation between the use of collocations and lexical bundles in improving the writing of first-year EFL college students. The findings confirmed that the use of collocations and lexical bundles positively impacted the quality of students' writing. However, it is difficult to confirm that correlation means causation. Improving writing is a complex and lengthy process. However, the use of collocation and lexical bundles was one factor that improved writing quality.

Thus, instead of questioning whether lexical bundles help improve EFL learners' writing skills, more research should be directed to how teachers and program administrators can develop more teaching materials to help EFL teachers use lexical bundles and collocations in writing classes. Scholars must focus more on developing classroom materials to teach students to use lexical bundles and collocations to improve their sentence structure. In this way, instead of writing short, simple sentences using lexical bundles, whether discourse, referential, or stance bundles, the students can develop their ability to compose more advanced and well-developed structured sentences.

To conclude, the study investigated the correlation between the explicit teaching of collocation and lexical bundles and the improvement of the writing skill of first-year university students. The results indicate a set of implications: There is no doubt that the explicit teaching of collocations and lexical bundles scaffolds the learning process and enhances the acquisition of the collocations and lexical bundles, which leads to the improvement of writing skills. The results also confirmed that adopting the communicative, interactive approach had significant results that were apparent in improving the writing skill of the experimental group.

The results also confirmed that the use of the attitudinal collocation and lexical bundles significantly improved the writing of the experimental group. The study's findings should encourage educators and administrators in foreign language programmes in universities to incorporate more collocations and lexical bundles in their syllabi. Book developers should design more teaching materials that help new teachers use collocations and lexical bundles to plan their language lessons. Educators should dedicate more research into the relationship between the use of collocations and lexical bundles and syntactic development because the use of collocations and lexical bundles will benefit EFL learners in learning how to develop their sentences and write more complex sentences that would eventually assist them in communicating their ideas more efficiently

5.2 Limitations

This study had some limitations, suggesting that the results should be dealt with carefully. First, the study's results depended very much on the participants' language level, which was significant in the results. If the study is repeated with different participants and in a different context, the results may be different, which means that the results of this study cannot be generalized. Second, this study had the risk of researcher bias, as I was the one who taught the classes. Researcher bias can affect any research point, from the study design or data collection to data analysis and reporting of the results. The disadvantage of researcher bias is that errors can occur during sampling or testing.

I took all the required measures to reduce researcher bias in my study. The results of the writing essays were determined by a committee of three English instructors who would read the essays and determine the overall grade in the pre- and post-intervention writing assessments. I wanted to train one of the teachers to avoid researcher bias, but this was not possible due to the regulations at my place of work. The university where I teach prohibits anyone from working for a short period.

5.3 Recommendations

The recommendation section is significant in any research because it serves as a road map for future research. Besides finding answers to existing research questions, good research should open the door to more studies; in other words, the results of one research should serve as the foundation or the starting point of more research. The current study investigated the correlation between the teaching of collocations and lexical bundles and the improvement of the writing skill of First-year university students. Considering the present study's findings, I will discuss a list of recommendations that would be useful for fellow researchers who plan to build on the current research findings. I will provide recommendations for future research regarding methodologies, design, and participants.

The current study investigated the correlation between teaching collocations and lexical bundles in improving the writing skill of first-year university students. The study chose to focus on writing skills because it took place in the English Language Bridge program at one of the universities, and the curriculum of the class was only writing. However, it would be very beneficial if the same study is repeated to investigate the impact of teaching collocations and lexical bundles on the four language skills, not just writing skills. When analyzing the effects of any teaching material, the researcher must test the listening skills, speaking, reading, and writing. However, the focus on the four skills was not possible in the current study because the learning outcomes of the syllabus bound me, and it was a writing course; besides, I was in no position to change the curriculum or the learning outcome of the course. I recommend that the study be repeated, targeting the four language skills.

The current study adopted a purely scientific approach that depends on the analysis of pre-tests and post-tests along with the results of the writing grade to investigate the effectiveness of the teaching materials that focus on collocation and lexical bundles. Nonetheless, I recommend adding an interview at the end of the intervention programme to listen to the participants and get their feedback on the teaching materials and whether the collocations and lexical bundles assisted them in improving their language skills.

Another recommendation is that a similar study is conducted on different students' levels to investigate how we can develop suitable materials for each EFL language level student, specifically in writing skills. I recommend contacting the Ministry of Education or any governmental authority with access to many students in schools or language programs.

The current study had 65 participants. To investigate the effectiveness of any language materials, any researcher needs a more significant number of students, not only many students but a longer duration of time. I conducted the study over 15 weeks which is the semester duration. I recommend that the study be conducted again on the students when they move to the next level to investigate if the information the participants learned in the course turned into acquired knowledge, not temporary. The participants will continue to use the collocations and lexical bundles in their writing. It is beneficial to run the study on many participants and over an extended period. We need to monitor student progress at different levels. Suppose the study began with students at the beginner level, intermediate, upper intermediate, until the advanced level. In that case, we could monitor their progress until they finish the program to investigate how they use their knowledge of collocations and lexical bundles.

5.4 Suggestions for Further Research

One suggestion is that scholars stop questioning whether collocations and lexical bundles benefit EFL learners' language development. While researching this thesis, I found many studies that asked the same questions. However, we must acknowledge that they are beneficial and dedicate more time and research to developing helpful teaching practices for teachers to incorporate collocations and lexical bundles into their classes.

Another recommendation is that researchers need to develop more research on which collocations and lexical bundles are more beneficial than others; besides, we need to investigate linking lexical bundles to syntactic development. The following research should be more directed at how we can use specific collocations and lexical bundles to assist EFL learners in knowing how to write more complex sentences with multiple clauses. The lexical bundles could be very beneficial when we introduce the various types of sentence structures, especially

complex and complex compound types of sentences—mastering the collocations and lexical bundles can help the EFL learners develop their sentence structure.

More research should be directed to syntactic complexity because it is a clear indicator of the language proficiency of EFL learners. Clearly, syntactic complexity requires more proficient users of the language; however, if we start by introducing the lexical bundles in beginners and intermediate classes and train the learners to use them to build complex sentences would have a positive impact on the EFL language proficiency in the long run.

5.5 Personal Reflection on the Thesis Journey

My Doctorate in Education journey has been one of the most significant experiences in my life. During the 7-year journey, I learned to be a critical thinker, especially when reading different publications and responding to the information in them.

It took me a while to move from being a teacher to a researcher. Now, I am more concerned with how teaching theories are generated, whereas earlier, I would apply any theory without questioning it.

I am now more concerned about helping my fellow English teachers in their practices, which I can do by publishing to spread knowledge.

I am glad I completed two publications during my EdD Programme: Kondos, S. (2018). The Effect of the Use of Technology on the Nature of Teacher's Profession. Arab World English Journal, 9(1) and Kondos, S. (2020). The Impact of Teaching Formulaic Sequences on Improving the Writing Skills of Freshman Students at an American University in the Middle East. *Journal of Higher Education Theory and Practice*, *20*(11).

I cherished the advice I received during the first year of my doctoral studies; to be a researcher is to share your findings for the betterment of the practice. Doctoral studies are about spreading knowledge to help others. It is only by sharing our research findings that we learn to be better scholars. I am glad I found my niche in formulaic language, especially collocations and lexical bundles. This EdD is only the beginning of my studies.

REFERENCES

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*(4), 235-247.

Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, *31*, 81–92.

Adolphs, S., & Schmitt, N. (2004). Vocabulary coverage according to spoken discourse context. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 39–49). John Benjamins.

Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. *Automatic Treatment and Analysis of Learner Corpus Data*, 249–264.

Al Hassan, L., & Wood, D. (2015). The effectiveness of focused instruction of formulaic sequences in augmenting L2 learners' academic writing skills: A quantitative research study. *Journal of English for Academic Purposes*, *17*, 51–62.

Alonso Ramos, M., Wanner, L., Vincze, O., Casamayor del Bosque, G., VázquezVeiga, N., Mosqueira Suárez, E., & Prieto González, S. (2010). Towards a motivated annotation schema of collocation errors in learner corpora. In *Proceedings of the International Conference on Language Resources and Evaluation* (pp. 3209–3214), May 17–23, 2010, Valletta, Malta.

Anastasi, A., & Urbina, S. (1997). *Psychological testing*. Prentice Hall/Pearson Education.

Appel, R., & Wood, D. (2016). Recurrent word combinations in EAP test-taker writing: Differences between high- and low-proficiency levels. *Language Assessment Quarterly*, *13*(1), 55–71.

Árabe, C. (2010). Teaching collocation in the Arabic as a foreign language class. *Enseñanza y aprendizaje de la lengua árabe*, 163.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Baker, P. (2006). *Using corpora in discourse analysis*. Continuum.

Bardakçı, M. (2016). Breadth and depth of vocabulary knowledge and their effects on L2 vocabulary profiles. *English Language Teaching*, *9*(4), 239–250. http://dx.doi.org/10.5539/elt.v9n4p239

Bestgen, Y., & Granger, S. (2014). Quantifying the development of phraseological competence in L2 English writing: An automated approach. *Journal of Second Language Writing*, *26*, 28–41.

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, *26*(3), 263–286.

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. 71–93). Peter Lang.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at…: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*(3), 371–405.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and us*e. Cambridge University Press.

Biber, D., & Gray, B. (2013). Discourse characteristics of writing and speaking task types on the TOEFL iBT test: A lexico-grammatical analysis. *TOEFL iBT Research Report*, *19*.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education.

Biesta, G. (2010). Pragmatism and the philosophical foundations of mixed methods research. *Sage Handbook of Mixed Methods in Social and Behavioral Research*, *2*, 95–118.

Birks, M., & Mills, J. (2015). *Grounded theory: A practical guide*. SAGE Publications, Inc.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language teaching research*, *10*(3), 245-261.

Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics*, *32*, 83-110.

Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the effects of exercises on verb–noun collocations. *Language Teaching Research*, *18*(1), 54-74.

Boers, F., Dang, T. C. T., & Strong, B. (2017). Comparing the effectiveness of phrase-focused exercises: A partial replication of Boers, Demecheleer, Coxhead, and Webb (2014). *Language Teaching Research*, *21*(3), 362-380.

Boisvert, R. D. (1998). *John Dewey: Rethinking our time*. SUNY Press.

Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research*, *1*(1), 8–22.

Byrd, P., & Coxhead, A. (2010). On the other hand: Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney Papers in TESOL*, *5*(5), 31–64.

Cadierno, T., & Robinson, P. (2009). Language typology, task complexity and the development of L2 lexicalization patterns for describing motions events. *Annual Review of Cognitive Linguistics*, *6*, 247–277.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of personality and social psychology*, *67*(2), 319.

Campbell, D. T., & Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research*. Ravenio Books.

Charmaz, K. (2000). Grounded theory: Objectivist and constructivist methods. *Handbook of Qualitative Research*, *2*, 509–535.

Chen, X. (2009). *The role of formulaic sequence in speech fluency development in ESL*. Paper presented at the International Conference on Computational Intelligence and Software Engineering, Wuhan, China.

Chen, W. C. (2012, November). Formulaic Language Use in Virtual Academic Discussions. In *Global Learn* (pp. 152-157). Association for the Advancement of Advancement of Computing in Education (AACE).

Chen, Y.-H., & Baker, P. (2016). Investigating criterial discourse features across second language development: Lexical bundles in rated learner essays, CEFR B1, B2 and C1. *Applied Linguistics*, *37*(6), 849–880.

Cobb, T. (1999). Breadth and depth of lexical acquisition with hands-on concordancing. *Computer Assisted Language Learning*, *12*(4), 345–360. http://dx.doi.org/10.1076/call.12.4.345.5699

Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education*. Routledge.

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.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, *23*, 397–423.

Cortes, V. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education*, *17*, 391–406.

Cowie, A. P. (1998). Introduction. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 1–20). Oxford University Press.

Creswell, J. W. (2008). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (3rd ed.). Pearson Prentice Hall.

Creswell, J. W. (2009). *Research design: Qualitative and mixed methods approaches*. SAGE Publications, Inc.

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. SAGE Publications, Inc.

Crossley, S., & Salsbury, T. L. (2011). The development of lexical bundle accuracy and pro- duction in English second language speakers. *IRAL*, *49*, 1–26.

Crotty, M. (1989). *The foundations of social research*. SAGE Publications, Inc.

Davis, M., & Morley, J. (2015). Phrasal intertextuality: The responses of academics from different disciplines to students' re-use of phrases. *Journal of Second Language Writing*, *28*, 20–35.

De Cock, S. (2000). Repetitive phrasal chunkiness and advanced EFL speech and writing. In C. Mair & M. Hundt (Eds.), *Corpus linguistics and linguistic theory* (pp. 51–68). Rodopi.

De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literature*, *2*, 225–246.

Deen, P. (2011). Herbert Marcuse's review of John Dewey's logic: The theory of inquiry. *Transactions of the Charles S. Peirce Society*, *46*(2), 258–265.

Denscombe, M. (2008). Communities of practice: A research paradigm for the mixed methods approach. *Journal of Mixed Methods Research*, *2*(3), 270–283.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2011). *The SAGE Handbook of Qualitative Research*. SAGE Publications, Inc.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process* (Vol. 8).

Dewey, J. (1938). *Logic: The theory of inquiry*. Henry Holt and Company.

Dewey, J. (1954). *The public and its problems*. Swallow Press.

Dontcheva-Navratilova, O. (2012). Lexical bundles in academic texts by non-native speakers.

Duemer, L. S., & Zebidi, A. (2009). The pragmatic paradigm: An epistemological framework for mixed methods research. *Journal of Philosophy and History of Education*, *59*(5).

Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics*, *38*(2), 165–193.

Durrant, P. (2018). Formulaic language in English for academic purposes. *Understanding Formulaic Language*, 211-227.

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

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

El-Dakhs, D. A. S., Prue, T. T., & Ijaz, A. (2017). The effect of the explicit instruction of formulaic sequences in pre-writing vocabulary activities on foreign language writing. *International Journal of Applied Linguistics and English Literature*, *6*(4), 21-31.

Ellis, N. C. (2008). Usage-based and form-focused language acquisition. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 372– 405). Routledge.

Ernest, P. (1994). *An introduction to research methodology and paradigms*. Hilary A. Radnor.

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

Frankel Pratt, S. (2016). Pragmatism as ontology, not (just) epistemology: Exploring the full horizon of pragmatism as an approach to IR theory. *International Studies Review*, *18*(3), 508–527.

Garnier, M., & Schmitt, N. (2015). The PHaVE List: A pedagogical list of phrasal verbs and their most frequent meaning senses. *Language Teaching Research*, *19*(6), 645–666.

Gass, S. (with J. Behney, & L. Plonsky). (2013). *Second language acquisition: An introductory course* (4th ed.). Routledge.

Gibbs, G. R. (2018). *Analyzing qualitative data* (Vol. 6). SAGE Publications, Inc.

Gibbsons, P. (2009). *English learners, academic literacy and thinking: Learning in the challenge zone.* Heinemann.

Gilquin, G. (2007). To err is not all: What corpus and elicitation can reveal about the use of collocations by learners. *Zeitschrift für Anglistik und Amerikanistik*, *55*(3), 273–291.

Gilquin, G., Granger, S., & Paquot, M. (2007). Learner corpora: The missing link in EAP pedagogy. *Journal of English for Academic Purposes*, *6*(4), 319–335.

Gilquin, G., & Gries, S. T. (2009). Corpora and experimental methods: A state-of-the-art review. *Corpus Linguistics and Linguistic Theory*, *5*(1), 1–26.

Goldkuhl, G. (2012). Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems*, *21*(2), 135–146.

Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, *11*(4), 341–363.

Graham, S., Richards, B. J., & Malvern, D. D. (2008). Progress in learning French vocabulary in a one year advanced course at school. *French Language Studies*, *18*, 394–364.

Granger, S. (2017). Academic phraseology: A key ingredient in successful L2 academic literacy. In R. Vatvedt Fjeld, K. Hagen, B. Henriksen, S. Johansson, S. Olsen, & J. Prentice (Eds.), *Academic language in a Nordic setting: Linguistic and educational perspectives* (Vol. 9(3), pp. 9–27). Studies in Language.

Granger, S. (2018). Formulaic sequences in learner corpora: Collocations and lexical bundles. In A. Siyanova-Chanturia & A. Pellicer-Sanchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 228–247). Routledge.

Grant, L., & Nation, P. (2006). How many idioms are there in English? *ITL International Journal of Applied Linguistics*, *15*(1), 1–14.

Greene, J., & Caracelli, V. (2003). Making paradigmatic sense of mixed methods practice. In A. Tashakkori & C. Teddlie, *Handbook of mixed methods in behavioral and social research.*

Grix, J. (2004). *The foundations of research*. Palgrave Macmillan.

Gyllstad, H. (2007). *Testing English collocations* [Unpublished dissertation]. Lund University.

Gyllstad, H. (2009). Designing and evaluating tests of receptive collocation knowledge: COLLEX and COLLMATCH. In *Researching collocations in another language* (pp. 153-170). Palgrave Macmillan, London.

Gyllstad, H., & Schmitt, N. (2019). Testing formulaic language. *Understanding formulaic language: A second language acquisition perspective*, 174–191.

Haastrup, K., & Henriksen, B. (2000). Vocabulary acquisition: Acquiring depth of knowledge through network building. *International Journal of Applied Linguistics*, *10*(2), 221-240.

Harmer, J. (2007). *The practice of English language teaching*. Longman.

Hatami, S., & Tavakoli, M. (2013). The role of depth versus breadth of vocabulary knowledge in success and ease in L2 lexical inferencing. *TESL Canada Journal*, *30*(1), 1.

Hazenberg, S., & Hulstun, J. H. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics*, *17*(2), 145–163. http://dx.doi.org/10.1093/applin/17.2.145

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, *21*, 30–317.

Heyworth, F. (2006). The common European framework. *ELT Journal*, *60*(2), 181-183.

Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation* (pp. 47–67). English Teaching Publications.

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

Howell, K. (2013). *An introduction to the philosophy of methodology*. SAGE Publications, Inc.

Hu, M., & Nation, I. S. P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language*, *23*(1), 403–430.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(1), 4–21.

Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics*, *32*, 150–169.

Jackendoff, R. (1995). The boundaries of the lexicon. In M. Everaert, E. van der Linden, A. Schenk, & R. Schreuder (Eds.), *Idioms: Structural and psychological perspectives* (pp. 133– 166). Lawrence Erlbaum.

Jones, M. A., & Haywood, S. (2004). Facilitating the acquisition of formulaic sequences: An exploratory study in an EAP context.

Kaushik, V., & Walsh, C. A. (2019). Pragmatism as a research paradigm and its implications for social work research. *Social Sciences*, *8*(9), 255.

Kerlinger, F. N. (1966). *Foundations of behavioral research*.

Kim, Y. (2009). The effects of task complexity on learner-learner interaction. *System*, *37*, 254–268.

Kondos, S. (2020). The impact of teaching formulaic sequences on improving the writing skills of freshman students at an American university in the Middle East. *Journal of Higher Education Theory & Practice*, *20*(11).

Kormos, J. (2006). *Speech production and second language acquisition*. Lawrence Erlbaum.

Kurniawan, I. (2016). Measuring EFL students vocabulary size: Why and how. *English Education: Jurnal Tadris Bahasa Inggris*, *9*(1), 89–102.

Laufer, B. (1989). A factor of difficulty in vocabulary learning: Deceptive transparency. *AILA review*, *6*(1), 10-20.

Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing*, *21*, 202–226.

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

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, *16*(1), 33–51.

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, *61*, 647–672.

Lewis, M. (1997). *Implementing the lexical approach: Putting theories into practice*. Language Teaching Publications.

Lewis, M. (2000). Teaching collocation: Further development in the lexical approach. Language Teaching Publications.

Liu, D. (2003). The most frequently used spoken American English idioms: A corpus analysis and its implications. *Tesol Quarterly*, *37*(4), 671-700.

Long, M., & Crook, G. (1992). *Three approaches to task-based syllabus design. TESOL Quarterly*, *26*, 27–55.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, *15*(4), 474–496.

Mackey, A., & Grass, S. (Eds.). (2006). *Interaction research: Extending the methodological boundaries.* Studies in Second Language Acquisition.

Mackenzie, N., & Knipe, S. (2006). Research paradigms, methods and methodology. *Issues in Educational Research*, *16*(2), 193–205.

Martinez, R. (2011). *Putting a test of multiword expressions to a test*. Paper presented at the IATEFL Testing, Evaluation and Assessment SIG. University of Innsbruck, September 16, 2011. https://ufpr.academia.edu/RonMartinez/Talks

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, *33*(3), 299–320.

Maxcy, S. J. (2003). Pragmatic threads in mixed methods research in the social sciences: The search for multiple modes of inquiry and the end of the philosophy of formalism. *Handbook of Mixed Methods in Social and Behavioral Research* (pp. 51–89).

McDonough, J., & McDonoughs, S. (1997). *Research methods for English language teachers*. Arnold.

McGilvary, E. B., Cunningham, G. W., Lewis, C. I., & Nagel, E. (1939). A symposium of reviews of John Dewey's logic: The theory of inquiry. *The Journal of Philosophy*, *36*(21), 561–581.

Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjær, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35–52). Cambridge University Press.

Meara, P. (1997). Towards a new approach to modelling vocabulary acquisition. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 109–121). Cambridge University Press.

Meara, P. (2002). The rediscovery of vocabulary. *Second Language Research*, *18*(4), 393–407. http://dx.doi.org/10.1191/0267658302sr211xx

Meara, P., & Rodriguez Sánchez, I. (1993). *Matrix models of vocabulary acquisition: An empirical assessment*. CREAL Symposium on Vocabulary Research.

Mertens, D. M. (2005). *Research methods in education and psychology: Integrating diversity with quantitative and qualitative approaches* (2nd ed.). SAGE Publications, Inc.

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

Miralpeix, I., & Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *International Review of Applied Linguistics in Language Teaching*, *56*(1), 1–24.

Moghaddam, A. (2006). Coding issues in grounded theory. *Issues in Educational Research*, *16*(1), 52–66.

Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, *1*(1), 48–76.

Morgan, D. L. (2014). Pragmatism as a paradigm for mixed methods research. *Integrating qualitative and quantitative methods: A pragmatic approach.* SAGE Publications, Inc.

Morrow, S. L. (2005). Quality and trustworthiness in qualitative research in counseling psychology. *Journal of Counseling Psychology*, *52*(2), 250.

Murcia, C. (2001). *Teaching English as a second or foreign language.* Heinle and Heinle Thomsong Learning.

Mutlu, G. K., & Kaşlıoğlu, Ö. (2016). Vocabulary size and collocational knowledge of Turkish EFL learners/İngilizceyi Yabancı Dil Olarak Öğrenen Türk Öğrencilerin Kelime Dağarcığı ve Eşdizim Bilgisi.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading research quarterly*, 304-330.

Nagy, W. E., Anderson, R., Schommer, M., Scott, J. A., & Stallman, A. (1989). Morphological families in the internal lexicon. *Reading Research Quarterly*, *24*(3), 263–282.

Nation, I.S.,& Nation, I. S. P. (2001). *Learning vocabulary in another language* (Vol.10). Cambridge: Cambridge university press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review/ La Revue canadienne des langues vivantes*, *63*(1), 59–82. http://dx.doi.org/10.3138/cmlr.63.1.59

Nation, I. S. P. (2010). *Learning vocabulary in another language*. Cambridge University Press.

Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford University Press.

Neuman, W. L., & Kreuger, L. (2003). *Social work research methods: Qualitative and quantitative approaches*. Allyn and Bacon.

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

O'Donnell, M. B., Römer, U., & Ellis, N. C. (2013). The development of formulaic sequences in first and second language writing. *International Journal of Corpus Linguistics*, *18*(1), 83–108.

Ormerod, R. (2006). The history and ideas of pragmatism. *Journal of the Operational Research Society*, *57*(8), 892–909.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, *24*(4), 492–518.

Patton, M. Q. (2002). *Qualitative research and evaluation methods*. SAGE Publications, Inc.

Paquot, M. (2013). Lexical bundles and L1 transfer effects. *International Journal of Corpus Linguistics*, *18*(3), 391–417.

Paquot, M. (2017). The phraseological dimension in interlanguage complexity research. *Second Language Research*. Online view. https://doi.org/10.1177/0267658317694221

Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, *32*, 130–149.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–225). Longman.

Pellicer-Sánchez, A., & Boers, F. (2018). Pedagogical approaches to the teaching and learning of formulaic language. *Understanding formulaic language*, 153-173.

Peters, E. (2012). Learning German formulaic sequences: The effect of two attention-drawing techniques. *The Language Learning Journal*, *40*(1), 65-79.

Pignot-Shahov, V. (2012). Measuring L2 receptive and productive vocabulary knowledge. *Language Studies Working Papers*, *4*(1), 37–45.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, *10*, 355–371.

Read, J. (2005). Applying lexical statistics to the IELTS speaking test. *Research Notes*, *20*, 12–16.

Read, J., & Nation, P. (2004). Measurement of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 23–35). John Benjamins.

Revier, R. L. (2009). Evaluating a new test of whole English collocations. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language: Multiple interpretations* (pp. 125–138). New York, NY: Palgrave Macmillan.

Richard, J. C., & Renandya, W. A. (2002). Approaches to teaching. *Methodology in language teaching: An anthology of current practice*, 1–5.

Rolfe, G. (2006). Validity, trustworthiness and rigour: Quality and the idea of qualitative research. *Journal of Advanced Nursing*, *53*(3), 304–310.

Rorty, R. (2000). Phailed prophecies, glorious hopes. In *Philosophy and Social Hope*.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta analysis of research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133–164). John Benjamins.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329–363.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Press.

Schmitt, N. (2022). Norbert Schmitt's essential bookshelf: Formulaic language. *Language Teaching*, 1-12.

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). John Benjamins.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the vocabulary levels test. *Language Testing*, *18*(1), 55–88.

Schmitt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, *14*, 357–385.

Schonell, F., Meddleton, I., Shaw, B., Routh, M., Popham, D., Gill, G., … Stephens, C. (1956). *A study of the oral vocabulary of adults*. University of Queensland Press / University of London Press.

Shahov, V. P. (2012). Measuring L2 receptive and productive vocabulary knowledge. *University of Reading: Language Studies Working Papers*, *4*, 37–45.

Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, *31*(4), 487–512.

Sinclair, J. (1987). *Looking up: An account of the COBUILD project in lexical computing*. Collins Cobuild.

Siyanova-Chanturia, A., & Pellicer-Sanchez, A. (Eds.). (2018). *Understanding formulaic language: A second language acquisition perspective*. Routledge.

Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning*, *63*(1), 121–159.

Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes*, *12*, 214–225.

Stuhr, J. J. (1999). Pragmatism and classical American philosophy: Essential readings and interpretive essays.

Sürücü, L., & MASLAKÇI, A. (2020). Validity and reliability in quantitative research. *Business & Management Studies: An International Journal*, *8*(3), 2694-2726.

Sugiarti, T. R., Fitrianasari, N. I., & Sulistyorini, T. (2018). Lexical bundles in academic writing by undergraduate and graduate students of English Language Education Program. *Loquen: English Studies Journal*, *11*(2), 1–14.

Suter, L. E. (2005). Multiple methods: Research methods in education projects at NSF. *International Journal of Research & Method in Education*, *28*(2), 171–181.

Szudarski, P. (2012). Effects of meaning-and form-focused instruction on the acquisition of verb-noun collocations in L2 English. *Journal of Second Language Teaching & Research*, *1*(2), 3-37.

Tashakkori, A., & Creswell, J. W. (2007). Exploring the nature of research questions in mixed. *Journal of Mixed Methods Research*, *1*(3), 207–211.

Tashakkori, A., & Teddlie, C. (Eds.). (2010). *SAGE Handbook of Mixed Methods in Social and Behavioral Research*. SAGE Publications, Inc.

Tashakkori, A., Teddlie, C., & Teddlie, C. B. (1998). *Mixed methodology: Combining qualitative and quantitative approaches* (Vol. 46). SAGE Publications, Inc.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. SAGE Publications, Inc.

Thornbury, S. (2002). *How to teach vocabulary*. Pearson Education Limited.

Tolley, E. E., Ulin, P. R., Mack, N., Robinson, E. T., & Succop, S. M. (2016). *Qualitative methods in public health: A field guide for applied research*. John Wiley & Sons.

Troudi, S. (2010). Pragmatic nature and theoretical framework in educational research. In M. Al-Hamly et al. (Eds.), *English in learning: Learning in English* (pp. 315–323). Dubai TESOL Arabia Publications.

Troudi, S., & Jendli, A. (2011). Emirati students' experiences of English as a medium instruction. In A. Al-Issa & L. Dahan (Eds.), *Global English and Arabic: Issues of language, culture, and identity* (pp. 23–48). Peter Lang.

Tsai, K.-J. (2015). Profiling the collocation use in ELT textbooks and learner writing. *Language Teaching Research*, *19*(6), 723–740.

Ünaldı, İ. (2011). *A comparative investigation of lexical networks of Turkish learners of English as a foreign language: A corpus based study* [Unpublished doctoral dissertation]. Çukurova University, Adana.

Vidakovic, I., & Barker, F. (2010). Use of words and multi-word units in skills for life writing examinations. *University of Cambridge ESOL Examinations Research Notes*, *41*, 7–14.

Vincze, O., García-Salido, M., Orol, A., & Alsonso-Ramos, M. (2016). A corpus study of Spanish as a foreign language learners' collocation production. In M. Alonso-Ramos (Ed.), *Spanish learner corpus research: Current trends and future perspectives* (pp. 299–331). John Benjamins.

Wang, Y. (2016). *The idiom principle and L1 influence: A contrastive learner-corpus study of delexical verb + noun collocations*. John Benjamins.

Weinert, R. (1995). The role of formulaic language in second language acquisition: A review. *Applied Linguistics*, *16*(2), 180–205.

Wood, D. (2002). Formulaic language acquisition and production: Implications for teaching. *TESL Canada Journal*, 01-15.

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

Yoon, H.-J. (2016). Association strength of verb-noun combinations in experienced NS and less experienced NNS writing: Longitudinal and cross-sectional findings. *Journal of Second Language Writing*, *34*, 42–57.

Yvonne Feilzer, M. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research*, *4*(1), 6–16.

APPENDIX 1: The List of the Collocations & Lexical Bundles

Group A: Referential Expressions

An attempt to	in accordance with	in the course of
Are/was based on	in such a way that	In the form of
Depend on	in the absence of	there are no
On the basis of	in terms of	there are several
With regard to	in this case of	in a number of
In terms of	degree to which	in some cases
Be related to	it has been	this does not
Which can be	does not have	this means that
Is more likely	on the other hand	the difference between
At the end of this point	at this stage	to distinguish between

Group B: Stance Expressions

Appear(s) to be	Are likely to	As a whole
Assumed to	Be argued that	Be explained by
Be regarded as	Been shown that	If they are
Is determined by	We have seen	Take in account
Can be achieved	Most likely to	Carried out by
Has been used	It should be noted	Take into account
Can be expressed	Can be achieved	Are able to

Group C: Discourse Organizing Functions

As shown in	Important role	It is necessary
It is obvious that	It is interesting	It is worth
It is difficult	In the present study	As a consequence,
As a result of	Due to the fact	Is affected
It follows	And if you	Even though
In conjunction	Due to the fact	To determine whether

APPENDIX 2: Assessing Collocations & Lexical Bundle Knowledge

Signposts to link two sentences:

On the other hand – Lastly – But on the other hand – For example – For instance – In other words – However – Consequently

Underline the signposts words in the following ideas, and complete the sentences in your own words. Say what direction the signpost words are signaling.

4. I am interested in history but----------
5. Studying abroad is worthwhile even though----------
6. Working in the library is uncomfortable----------
7. The internet has changed the way we all live because----------
8. The climate of South East Asia is tropical. By contrast ----------
9. The effects of global warming are evident everywhere. For example, ----------.
10. Learning a foreign language can be difficult and at times frustrating. However, ----------
11. My brother never studied much at school, and consequently----------

Exercise:

The following formulaic sequences are useful in helping you be more precise:

By this I (don't) mean – In other words – To be more precise – In fact – That is to say – That is not to say that

Use one of the above expressions to add another sentence to the following arguments.

1- In my opinion, extended families are more successful than nuclear families.
2- It is often said that young people are more tolerant than older people.
3- I am convinced that choice is more important in the school curriculum.
4- Job commitment is not always the key to success.

*Adopted from V. Williams & C. McDowell, *New insights into IELTS*, Cambridge University Press.

APPENDIX 3: Accessing Formulaic Sequences Knowledge (Cont'd)

Look at the formulaic sequences in the box. These are the target sequences for this week. Use the scale to score yourself on each formulaic sequence. After we finish this week, score yourself again to check your improvement.

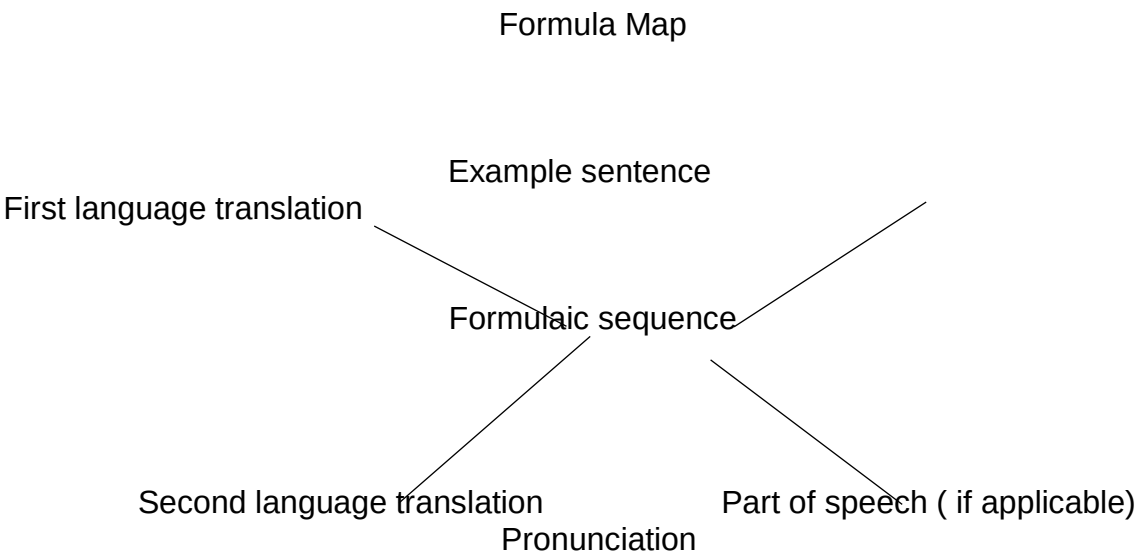
1- I don't know this formulaic sequence.
2- I have seen or heard this formulaic sequence, but I am not sure of the meaning.
3- I understand it when I see it or hear it in a sentence, but I don't know how to use it in my own writing
4- I know this formulaic sequence and can use it in my own writing

Appear(s) to be	Are likely to	As a whole
Assumed to	Be argued that	Be explained by
Be regarded as	Been shown that	If they are
Is determined by	We have seen	Take in account
Can be achieved	Most likely to	Carried out by
Has been used	It should be noted	Take into account
Can be expressed	Can be achieved	Are able to

*Adopted from D. Schmitt, N. Schmitt, & D. Mann, *Focus on vocabulary 1: Bridging vocabulary*, Pearson Longman

APPENDIX 4: Formula Map

Collocation & Lexical Bundle Card



*Adopted from D. Schmitt, N. Schmitt, & D. Mann, *Focus on vocabulary 1: Bridging vocabulary*, Pearson Longman

APPENDIX 5: Samples of In-Class Exercise Questions

Fill in the blanks with the following formula:

it follows, *as a result*, *even though*, *in this case of*, *appears to be*, *in terms of*, *be explained by*, *and if you*, *at this stage*, *there are several*, *it is difficult*

At this point, global warming …… one of the most prevalent issues on the rise in the world. Ever since the beginning of time, we have been diminishing the world's environments. Through our harming practices, we initiated global warming (also known as climate change) which can …….. such acts like burning fossil fuels, deforestation, and the emission of greenhouse gases. ……. of this, global warming seems to become a more vital issue that demands our recognition ……. As ……., governments around the world should be putting their share of effort in striving for change; ……. Solutions, …….. solutions, such as promoting less transportation by investing in more public transport or supporting the use of renewable energy sources. ……, undoubtedly, ……. to stop all our polluting practices, it only takes mere consciousness and awareness of how much you drive, eat meat, or just observing how much energy you consume; ……. do your best to be conscious of what you're doing and do your best to help what you can, then it is sufficient. ……. such a vital issue, all effort should be taken to prevent it, however minor or major the act may be.

Multiple Choice Questions:

Choose the right formulas from the below list:

1- It --------- how much he liked her
   made of – was obvious – on the other hand – even though
2- She could not tell ---------- between right and wrong.
   is more likely – the dark – be obvious – the difference
3- It should ---------- that all exams will be rescheduled.
   be noted – be there – be alive – be late
4- ---------- of the current teacher, all classes will be cancelled
   in the morning – after a while – take in account – in the absence of
5- It ---------- that the class is cancelled
   was ready – was assumed – was afraid – was dependent on

Appendix 6: Samples of In-class Reading Practice

After reading the passage, write the letter of the best answer to each question. 1 For as long as most of us can remember, shark has been synonymous with "danger." 2 Although the chances of being killed by a shark are one in 265 million, and ten times as many people are killed every year by random airplane parts falling out of the sky as are eaten by sharks, we still harbor an irrational fear of these creatures. 3 But it wasn't always this way.

4 As recently as 100 years ago, sharks were approached with no more trepidation than one might experience when encountering a jellyfish or a stingray. 5 Perhaps, on occasion, a beachgoer or fisherman might have an accidental run-in with a curious shark, and a newspaper story would mention injury or, very rarely, death as the result of what nearly everyone considered an inadvertent mishap. 6 But no one believed that sharks harbored some kind of innate hatred of humans and intentionally sought them out as victims—until the summer of 1916.

7 That summer, a young and successful businessman, Charles Vansant, decided to go for a quick swim one evening at a beach along the New Jersey shoreline. 8 Within minutes, he was screaming and thrashing in a pool of blood. 9 As rescuers attempted to pull him ashore, an unusually vicious shark remained clamped to Vansant's leg, releasing its hold only when Vansant was nearly on dry land. 10 Vansant died within hours. 11 Five more deadly shark attacks followed in quick succession at nearby beaches, leading the media to speculate that this was all the work of a single malicious shark. 12 Shark hysteria had begun. 13 Labeled as "killing machines," sharks were hunted and hated, and when dozens of American sailors were attacked by sharks after their ship sank in infested waters in the Pacific during World War II, the fervor reached new heights.

14 But perhaps the zenith of shark terror arrived when the film adaptation of Jaws was released in 1974. 15 Portraying the infamous great white shark as a ridiculously oversized and single-minded psycho killer seeking revenge, this movie and its sequels led to an all-out war on sharks. 16 By the 1980s, it was estimated that for every human being killed by a shark, two million sharks were slaughtered for sport, skin, meat, and fins. 17 What had been a thriving and mostly harmless sea creature was now perilously close to being endangered.

18 In 2000, Jaws author Peter Benchley expressed his regret for writing a book that had created so much fear and hatred, and he vowed to spend the rest of his life working to protect sharks. 19 "Today we know that almost every attack by sharks on humans is an accident," Benchley said. 20 "Considering all the knowledge gathered about sharks in the past 25 years, I couldn't possibly write Jaws today."

___ 1. In sentence 4, the word trepidation means
A. affection.
B. curiosity.
C. outrage.
D. fear.

___ 2. According to the passage, before 1916,

A. there were very few deaths from sharks.
B. sharks were less vicious than they are today.
C. newspaper stories seldom mentioned shark attacks.
D. people did not particularly fear sharks.

___ 3. According to the passage,
A. most people now realize that shark attacks are extremely rare.
B. sharks are mostly harmless.
C. people have about the same chance of being attacked by a shark as of being hit by falling airplane parts.
D. 1916 marked the first time anyone was attacked by a shark.

___ 4. In paragraphs 2–5, the author uses an organizational pattern that
A. lists several reasons why sharks attack humans.
B. shows, in time order, how our view of sharks has changed over the years.
C. contrasts great white sharks with other types of sharks.
D. presents the problem of shark attacks and offers a solution

___ 5. The relationship between the first half and the second half of sentence 11 is one of
A. cause and effect.
B. contrast.
C. illustration.
D. comparison.

___6. A conclusion that can be drawn from this passage is that
A. people are foolish to believe everything they read.
B. sharks are not dangerous.
C. sharks need our protection.
D. sharks will become extinct in a few years.

___ 7. The author suggests that
A. sharks are no more dangerous than any other sea creature.
B. sharks became more dangerous as more people began swimming in the ocean.
C. fear of sharks is an overreaction to media reports and unrealistic books and movies.
D. in a few years, people will no longer fear sharks.

___ 8. The tone of this passage is
A. informative.
B. indignant.
C. pessimistic.
D. pleading.

___ 9. Which sentence best expresses the central point of the passage?
A. The news media has a great effect on the shaping of public opinion.
B. Despite the movie Jaws and its sequels, we have nothing to fear from sharks.
C. Sharks were not always viewed as dangerous.
D. Our fear of sharks is irrational and harmful to sharks.

___ 10. Which statement offers the best support for the author's point that the word shark has become synonymous with "danger?"
A. "[T]he chances of being killed by a shark are one in 265 million."
B. "As recently as 100 years ago, sharks were approached with no more trepidation than one might experience when encountering a jellyfish or a stingray."
C. "Labeled as 'killing machines,' sharks were hunted and hated, and when dozens of American sailors were attacked by sharks after their ship sank in infested waters in the Pacific during World War II, the fervor reached new heights."
D. "In 2000, Jaws author Peter Benchley expressed his regret for writing a book that had created so much fear and hatred, and he vowed to spend the rest of his life working to protect sharks."

*Adapted from J. Langan (2013), *Ten steps to advanced reading as class text*, Townsend Press.

Appendix 7: Samples of In-Class Writing Practice

Complete the Model Essay by choosing the correct collocation & Lexical Bundles

President Richard Nixon used to keep an enemies list of all the people he didn't especially
like. I'm ashamed to confess it, but I, too, have an enemies list—a mental one. On this list are all the people I would gladly live without, the ones who cause my blood pressure to rise to the boiling point. The top three places on the list go to people with annoying nervous habits, people who talk in movie theaters, and people who talk on cell phones while driving

For example
First of all
Another
However
----------there are the people with annoying nervous habits.----------there are the ones who make faces. When in deep thought, they twitch, squint, and frown, and they can be a real distraction when I'm trying to concentrate during an exam.---------- type of nervous character makes useless designs. These people bend paper clips into
abstract sculptures or string the clips into necklaces as they talk.----------neither of these groups is as bad as the people who make noises. These individuals, when they are feeling uncomfortable, bite their fingernails or crack their knuckles. If they have a pencil in their hands, they tap it rhythmically against whatever surface is handy—a desk, a book, a head. Lacking a pencil to play with, they jingle the loose change or keys in their pockets. These people make me wish I were hard of hearing.

On the contrary
Then
As a result
After
Second

A ----------category of people I would gladly do away with is the ones who talk in movie theaters. These people are not content to sit back, relax, and enjoy the film they have paid to. ----------, they feel compelled to comment loudly on everything from the hero's hairstyle to the appropriateness of the background music.----------, no one hears a word of any dialogue except theirs. ----------these people have been in the theater for a while, their interest in the movie may fade. ---------- they will start discussing other things, and the people around them will be treated to an instant replay of the latest family scandal or soap-opera
episode. These stories may be entertaining, but they don't belong in a movie theater.

In addition
But
Last of all

----------, there are the people who talk on the phone while they're driving. One of the
things that irritates me about them is the way they seem to be showing off. They're saying,
“Look at me! I'm so important I have to make phone calls in my car.”----------, such
behavior is just plain dangerous. Instead of concentrating on adjusting carefully to everchanging
traffic conditions, they're weaving all over the road or getting much too close to the car in front of them as they gossip with a friend, make an appointment with a doctor, or order
a pizza.

So long as murder remains illegal, the nervous twitchers, movie talkers, and cell-phone users
of the world are safe from me.---------- if ever I am granted the power of life or death, these
people had better think twice about annoying me. They might not have long to live.

Editing Practice

Locate ten sentence-skills mistakes in the following passage. The mistakes are listed in the box below. As you locate each mistake, Write the number of the word group in the space provided. Then, in the space between the lines, edit and correct each mistake.

2 fragments ---------- ----------
2 missing lexical Bundles ---------- ----------
1 run- on ----------
2 mistakes in verb forms ---------- ----------
2 mistakes in collocations ---------- ----------
1 mistake modifier ----------

1The thirtieth -anniversary party of my uncle and aunt was the worst family gathering I have ever attended. 2 on a hot Saturday morning in July, Mom and I arrived out into the country to Uncle Ted's house . 3 It had already rained heavily , and the one place I left to park was in a muddy field. 4 Then, you could not believe the crowd. 5 There must have been two hundred people in Uncle Ted's small yard, including his five daughters with their husbands and children , all the other relatives , all the neighbors , and the entire congregation of their church.6 Since the ground was soaked with light rain was failing.7Mom and me went under the rented canopy with everybody else.8 we couldn't move between the tables, and the humidity fogged my glasses.9 After wiping my glasses, I seen that there is a lot of food . It was mainly cold chicken bd potato and macroni salads, I ate some lot because there was nothing else to do.11 We were surprised that Uncle Ted and his wife were doing all the work themselves. 12 They ran back and forth with trays of food and gathered trash into plastic bags staggering with exhaustion . 13 It is not seem like much of a way to celebrate . 14 Mom had upset that she is get to speak with them . 15 When we left I was hot , sticky to my stomach from overeating. 16 But quickly pushed our car out of the mud and got us on the road . 117 I have never is happier to leave a party.

Writing Thesis Statement Practice

Write a thesis for each group of supporting statements. This activity will give you practice in writing an effective essay thesis—one that is neither too broad nor too narrow. It will also help you understand the logical relationship between a thesis and its supporting details.

1. Thesis:----------
1-My first car was a rebellious-looking one that matched the way I felt and acted as a teenager.
2- My next car reflected my more mature and practical adult self.

3- My latest car seems to tell me that I'm aging; it shows my growing concern with comfort and safety.

2. Thesis: ----------

1. All the course credits that are accumulated can be transferred to a four-year school.
2. Going to a two-year college can save a great deal of money in tuition and other fees.
3. If the college is nearby, there are also significant savings in everyday living expenses.

3. Thesis: ----------

1. First, I tried simply avoiding the snacks aisle of the supermarket.
2. Then I started limiting myself to only five units of any given snack.
3. Finally, in desperation, I began keeping the cellophane bags of snacks in a padlocked cupboard.

ADOPTED FROM COLLEGE WRITING SKILLS WITH READINGS, SEVENTH EDITION INTERNATIONAL EDITION 2008

Editing Exercises (Cont'd)

Even though I love movies, my friends have stopped asking me to go. There are just too many problems involved in going to the movies. There are no small theaters anymore, I have to drive fifteen minutes to a big multaplex. Because of a supermarket and restarants, the parking lot is filled. I have to keep driving around to find a space. Then I have to stand in a long line. Hoping that they do not run out of tickets. Finally, I have to pay too much money for a ticket. Putting out that much money, I should not have to deal with a floor that seems coated with rubber cement. By the end of a movie, my shoes are often sealed to a mix of spilled soda, bubble gum, and other stuff.

The theater offers temptatons in the form of snacks I really don't need. Like most of us I have to worry about weight gain. At home I do pretty well by simply watching what I keep in the house and not buying stuff that is bad for me. I can make do with healthy snacks because there is nothing in the house. Going to the theater is like spending my evening in a 7- Eleven that's been equiped with a movie screen and there are seats which are comfortable. I try to persuade myself to just have a diet soda. The smell of popcorn soon overcomes me. My friends are as bad as I am. Choclate bars seem to jump into your hands, I am eating enormous mouthfuls of milk duds. By the time I leave the theater I feel sick and tired of myself. Some of the other moviegoers are the worst problem. There are teenagers who try to impress their friends in one way or another. Little kids race up and down the aisles, gigling and laughing. Adults act as if they're watching the movie at home. They talk loudly about the ages of the stars and give away the plot. Other people are droping popcorn tubs or cups of soda crushed ice and soda on the floor. Also coughing a lot and doing other stuff—bms! I decided one night that I was not going to be a moviegoer anymore. I joined Netflix, and I'll watch movies comfortable in my own living room.

1. Diane has a very ----------—no more than an opening sentence and a second sentence that states the ---------- She knows she can develop the introduction more fully in a later draft.

2. Of Diane's three supporting paragraphs, only the ----------paragraph lacks a topic sentence. She realizes that this is something to work on in the next draft.
3. There are some misspellings—for example, ---------- equiped . Diane doesn't worry about spelling at this point. She just wants to get down as much of the substance of her paper as possible.
4. There are various punctuation errors, such as the run-on sentences in the second----------and ---------paragraphs. Again, Diane is focusing on content; she knows she can attend to punctuation and grammar later. 5. At several points in the essay, Diane revises on the spot to make images more ---------- she changes “is sticky” to “seems coated with rubber cement,” “market” to “7-Eleven,” and “cups of soda” to “cups of crushed ice and soda.” 6. Near the end of her essay, Diane can't think of added details to insert so she simply puts the letters ---------- at that point to remind herself to “be more specific” in the next draft. She then goes on to finish her first draft. 7. Her ----------is as brief as her introduction. Diane knows she can round off her essay more fully during revision.

ADOPTED FROM COLLEGE WRITING SKILLS WITH READINGS, SEVENTH EDITION INTERNATIONAL EDITION 2008

APPENDIX 8: Sample of Collocations & lexical Bundles Tests

Test A

A. Fill in each blank with an appropriate transition from the box. Use each transition once. Then, in the space provided, write the letter of the transition you have chosen.

A. also B. final C. in addition D. later E. until

___ 1. One way to avoid tempting car thieves is by locking all valuables in the trunk or glove compartment. [2]You can _____ discourage thieves with tow trucks by parking in the middle of a block on a busy, well-lit street.

___ 2. The earliest humans probably used the lengthening and shortening of shadows on the ground to measure the passage of time. _____, the sundial was invented to tell time more precisely, but still by using the shadow principle.

___ 3. Hay fever is a genetic abnormality; it runs in families. [2]Some individuals simply are genetically predisposed to overreact to certain kinds of pollen. _____, infections, emotional stress, and changes in temperature may trigger immune reactions to pollen.

___ 4. The motion picture, invented in 1889, developed as an important form of entertainment during the first decade of the 1900s. [2]At first, the "nickelodeons," as the early movie theaters were called, appealed mainly to a lower-class and largely ethnic audience. [3]In 1902, New York City had 50 theaters; by 1908, there were more than 400 showing 30-minute dramas and romances. Not _____ World War I, when D.W. Griffith produced long feature films, did the movies begin to attract a middle-class audience.

___ 5. What causes people to join groups? One reason is for security, a factor that leads people to form neighborhood-watch groups. Another common reason for joining a group is a desire to be with others who share one's interests and values. Some people, for instance, join computer support groups to share ideas, knowledge, and software. Managers may join service groups, such as Rotary Clubs, to exchange ideas with other managers. Individuals may also form groups to acquire power that is difficult if not impossible to attain alone. Membership in a union or employee association, for example, provides workers with influence that they lack as individual employees. Goal accomplishment is a _____ reason people join groups. Mountain climbers and astronauts generally function in groups.

(Continues on next page)

B. (6–9.) Fill in each blank with an appropriate transition from the box. Use each transition once.

A. during B. first C. second D. when

During the (6)_____, or dilation, stage of the birth process, the uterus contracts, and the cervix flattens and dilates to allow the fetus to pass through. This labor stage can last from about two to sixteen hours, or even longer; it tends to be longer with the first child. (7)_____ the contractions start, they usually come at approximately fifteen- to twenty-minute intervals and are generally mild. Near the end of this first stage, the contractions change, becoming more difficult, longer, and more frequent. This period, lasting about an hour, is called "transition" and is the most difficult part of labor for many women. The (8)_____ stage of birth involves the actual delivery of the baby. This expulsion stage is quite variable and can last anywhere from two to sixty minutes or more. In the average delivery, the baby's head appears first, an event referred to as crowning. The rest of the body soon follows. The third stage of the birth process involves the delivery of the placenta (or afterbirth) and fetal membranes. (9)_____ this stage, mild contractions continue for some time. They help decrease the blood flow to the uterus and reduce the uterus to normal size.

___ 10. The pattern of organization of the above selection is
A. list of items.
B. time order.

Test B

A. Fill in each blank with an appropriate transition from the box. Use each transition once. Then, in the space provided, write the letter of the transition you have chosen.

A. during B. in addition C. last of all D. second E. when

___ 1. [1]Babies respond to the same four categories of tastes (sweet, sour, bitter, and salty) as adults do. [2]_____, they can identify familiar body odors and can even discriminate their mother's smell from the smell of another woman.

___ 2. [1]Sociologists have several basic methods of doing research. First is the experiment, which is useful for clearly defined questions in which varying factors can be controlled. A _____ method sociologists use is the survey, which is useful for gaining facts about a particular group; in order to be sound, the survey must be random. Direct observation is helpful for in-depth studies of social processes, but to be useful, such observations must be made by a skilled researcher. Finally, existing information can be studied as the basis for new conclusions.

___ 3. [1]Warren G. Harding was the first president elected after passage of the 19th Amendment, which gave women the right to vote. [2]_____ his presidency was rocked by scandals, some blamed women for supporting the candidate who photographed well and was considered handsome.

___ 4. [1]Throughout history, more than a few famous people were vegetarians. [2]One was Pythagoras, the famous Greek mathematician who lived in the fifth century B.C. and kept to a diet of bread, honey, and vegetables. [3]Plato and Aristotle were also vegetarians. [4]The poet Percy Bysshe Shelley became a convert to vegetarianism at twenty-one and later wrote a pamphlet in its defense. [5]_____, Count Tolstoy, the Russian author famous for *War and Peace*, not only ate no animals; he would also eat no eggs.

___ 5. [1]Although not generally considered a successful president, John Quincy Adams was certainly one of our most politically active ex-presidents. [2]After failing to win reelection in 1828, he won election to the House of Representatives in 1830. _____ his long tenure as congressman, he became, in the words of a Virginia congressman, "the acutest, the astutest, the archest enemy of Southern slavery that ever existed." [4]In 1841, Adams successfully argued before the Supreme Court that the Africans who had taken hold of the slave ship *Amistad*, on which they had been imprisoned, should be considered free and have the option to remain within the U.S. or return home as free people.

B. (6–9.) Fill in each blank with an appropriate transition from the box. Use each transition once.

A. during B. following C. immediately D. when

In the years (6)_____ World War II, William Levitt, a brassy New York developer, led the suburban revolution. [2]Born in Brooklyn in 1907, he dropped out of New York University because "I got itchy. [3]I wanted to make a lot of money. [4]I wanted a big car and a lot of clothes." [5]He got his wish. [6]Levitt and his brother made a fortune (7)_____ the depression by building houses. [7]But the Levitts really struck it rich after the war, (8)_____ the demand for new housing skyrocketed, and they developed an efficient system of mass production. [8]In 1947, on 1,200 flat acres of Long Island farmland, they built 10,600 houses that were (9)_____ sold and inhabited by more than 40,000 people—mostly young adults under thirty-five and their children. [9]The resulting community—the first mass-produced suburb in the United States—was named Levittown. [10]"Everyone is so young," one Levittowner noted, "that sometimes it's hard to remember how to get along with older people."

___ 10. The pattern of organization of the above selection is
A. list of items.
B. time order.

(continues on next page)

TEST C

A. Fill in each blank with an appropriate transition from the box. Use each transition once. Then, in the space provided, write the letter of the transition you have chosen.

A. although B. for instance C. however D. likewise E. so

___ 1. [1]In the very first nationally televised presidential debates, during the 1960 campaign, John F. Kennedy was tanned, confident, and dynamic, while Richard M. Nixon, recently ill, appeared drawn, haggard, and hesitant. _____ radio listeners judged the contest a draw, the TV audience gave the edge to Kennedy, who shot up irreversibly in the polls.

___ 2. [1]Children of homosexuals are no more likely to be homosexual themselves than are children of heterosexuals. [2]In one study, the vast majority of adult sons of gay fathers were heterosexual. [3]_____, a study of adult children of lesbians found that a large majority identified themselves as heterosexual.

___ 3. [1]In the late 1400s, Queen Isabella of Spain decreed slavery to be illegal unless the people involved were so wicked that their conditions as slaves would be better than as free men. [2]This legal requirement probably led Spanish conquerors to exaggerate the extent of cannibalistic practices among the native peoples they encountered _____ that they could enslave them.

___ 4. [1]In early America, many men looked upon taverns as refuges where they could drink, gamble, share stories, complain about bosses, and even organize over workplace issues. _____, many employers, Protestant clergymen, and female temperance reformers looked upon taverns as “the devil’s playground.”

___ 5. [1]Advertisers seek to determine an audience’s characteristics. [2]The analysis of observable audience characteristics is called demographics. [3]Demographics are composed of data about a target audience’s sex, age, income level, marital status, geographic location and occupation. [4]These data are observable because they are available to advertising agencies through census data and other sources. [5]Advertising agencies use demographic audience analysis to help advertisers target their messages. [6]_____, a motorcycle dealer certainly wouldn’t want to advertise in a baby magazine; a candy manufacturer probably wouldn’t profit from advertising in a diet and exercise magazine. [7]Advertising agencies try to match a client’s product to a thoroughly defined audience so each advertising dollar is well spent.

(continues on next page)

TEST D

Read each textbook paragraph below. Then answer the questions that follow.

A. [1]For years, various sociologists have struggled to determine the causes of poverty in America. [2]One approach holds that the poor are primarily responsible for their own poverty. [3]Throughout our history, people in the United States have valued self-reliance, convinced that social standing is mostly a matter of individual talent and effort. [4]This view sees society as offering plenty of opportunity to anyone able and willing to take advantage of it. [5]From this point of view, the poor are those who cannot or will not work, women and men with fewer skills, less schooling, and little motivation. [6]An alternative position holds that society is primarily responsible for poverty. [7]Sociologists who hold this view point to the loss of jobs in our inner cities as the primary cause of poverty, claiming there is simply not enough work to support families. [8]Thus, any apparent lack of trying on the part of the poor is a result of little opportunity rather than a cause of poverty.

___ 1. The main pattern of organization of the paragraph is
A. definition and example. B. cause and effect.
C. comparison. D. contrast.

2. One transition that signals the pattern of organization of this paragraph is _____.

B. [1]Words have two kinds of meanings—denotative and connotative. [2]Denotative meaning is precise, literal, and objective. [3]It simply describes the object, person, place, idea, or event to which the word refers. [4]One way to think of a word's denotative meaning is as its dictionary definition. [5]For example, denotatively, the noun *school* means "a place, institution, or building where instruction is given." [6]Connotative meaning is more variable, figurative, and subjective. [7]Put simply, the connotative meaning is what the word suggests or implies. [8]For instance, the connotative meaning of the word *school* includes the feelings, associations, and emotions that the word touches off in different people. [9]For some people, *school* might connote personal growth, childhood friends, and a special teacher. [10]For others, it might connote frustration, discipline, and boring homework assignments.

___ 3. The main pattern of organization of the paragraph is
A. definition and example. B. cause and effect.
C. comparison. D. contrast.

4. One transition that signals the pattern of organization of this paragraph is _____.

C. [1]The biological significance of a single sperm is very different from that of a single egg. [2]For healthy men, sperm is a "renewable resource" produced by the testes throughout most of the life course. [3]A man releases hundreds of millions of sperm in a single ejaculation, technically enough to fertilize every woman in North America. [4]However, a newborn girl's ovaries contain her entire lifetime allotment of immature eggs. [5]A woman releases a single egg from the ovaries every month. [6]Whereas men are biologically capable of fathering thousands of offspring, a woman is able to bear a much smaller number of children.

___ 5. The main pattern of organization of the paragraph is
A. definition and example. B. cause and effect.
C. comparison. D. contrast.

6. One transition that signals the pattern of organization of this paragraph is _____.

D. [1]The Roman (Western) alphabet consists of twenty-six letters, each representing a different sound. [2]Chinese writing incorporates about thirty thousand pictograms, each representing a different word. [3]Roman letters can be combined to produce any word; Chinese pictograms can be combined to express any idea. [4]Obviously, one advantage of the Roman system is that only a small number of symbols is required. [5]The twenty-six letters are easy to learn. [6]By contrast, the thirty thousand Chinese pictograms take a long time to learn. [7]And it is difficult to use them with a keyboard. [8]But when the Roman alphabet is used, only people who speak a particular language can read text written in that language. [9]On the other hand, any language can be written and read in Chinese pictograms, as long as the reader has the "key" to those pictograms. [10]In fact, speakers of many Chinese dialects who cannot understand one another's spoken language are able to communicate easily through text written in Chinese.

___ 7. The main pattern of organization of the paragraph is
A. definition and example. B. cause and effect.
C. comparison. D. contrast.

8. One transition that signals the pattern of organization of this paragraph is _____.

Test E

Read each textbook paragraph below. Then answer the questions that follow.

A. [1]Often governmental efforts to manipulate public opinion backfire when the public is made aware of the government's tactics. [2]Thus, in 1971, the United States government's attempts to build popular support for the Vietnam War were hurt when CBS News aired its documentary *The Selling of the Pentagon*, which revealed the extent and character of government efforts to sway popular sentiment. [3]In this documentary, CBS demonstrated the techniques, including planted news stories and faked film footage, that the government had used to misrepresent its activities in Vietnam. [4]These revelations, of course, had the effect of undermining public trust in all government claims.

___ 1. The main pattern of organization of the paragraph is
A. definition and example. B. cause and effect.
C. comparison. D. contrast.

2. One transition that signals the pattern of organization of this paragraph is _____.

B. [1]Both our 17th president, Andrew Johnson, and our 42nd president, Bill Clinton, were raised by a hard-working single mother. [2]Both were Southern Democrats. [3]Both were impeached. [4]Yet these superficial similarities in the careers of the two men mask dramatic differences between them. [5]When Andrew Johnson fired Secretary of War Stanton, he was charged with violating the Tenure of Office Act, a law that probably was unconstitutional. [6]His impeachment took place against the backdrop of a national debate over the course of post-Civil War Reconstruction, a debate that was intensely ideological. [7]In contrast, Bill Clinton was charged with lying and obstructing justice, charges that stemmed from an extramarital affair the president had conducted with a young White House intern, Monica Lewinsky. [8]Clinton's impeachment reflected the bitter partisanship and personal animosity between Democrats and Republicans in the 1990s.

___ 3. The main pattern of organization of the paragraph is
A. definition and example. B. cause and effect.
C. comparison. D. contrast.

4. One transition that signals the pattern of organization of this paragraph is _____

C. [1]Almost all effective interpersonal communication requires some degree of self-disclosure. [2]The very process of making friends involves learning more about each other. [3]In the broadest sense, self-disclosure means sharing biographical data, personal ideas, and feelings. [4]Statements such as "I was 5′6″ in seventh grade" reveal biographical information—facts about you as an individual. [5]Statements such as "I don't think prisons ever really rehabilitate criminals" disclose personal ideas and reveal what and how you think. [6]Statements such as "I get scared when I have to make a speech" disclose feelings. [7]Biographical disclosures are easiest to make, for they are, in a manner of speaking, a matter of public record. [8]It is statements about personal ideas and feelings that most people think of as self-disclosure.

___ 5. The main pattern of organization of the paragraph is
A. definition and example. B. cause and effect.
C. comparison. D. contrast.

6. One transition that signals the pattern of organization of this paragraph is _____.

D. [1]The rubber balloon is an interesting elastic object. [2]As it's inflated with helium, the rubber balloon is stretched away from its equilibrium shape, and it exerts restoring forces that try to return it to that equilibrium shape. [3]As a result, each region of the balloon's surface experiences three forces: an inward force from the pressure of air outside, an outward force from the pressure of helium inside, and an inward force from the elastic skin of the balloon itself. [4]Since each region of surface is stationary, it must be experiencing zero net force; the outward force must balance the two inward forces. [5]Therefore, the pressure of the helium inside the balloon must be somewhat greater than the pressure of the outside air.

___ 7. The main pattern of organization of the paragraph is

*Adapted from J. Langan (2013), *Ten steps to advanced reading as class text*, Townsend Press.

Appendix 9: Writing Assessment worksheet

Directions: Write an **essay** of **750-850** words on one of the topics below. The outline should include the following:

- a thesis statement,
- 3 topic sentences,
- extended examples

Your final version must include an introduction and a conclusion, with 3 body paragraphs.

TOPIC CHOICES

1- People work because they need money to live. What are some **other** reasons that people work? Discuss one or more of these reasons. Use specific examples and details to support your answer.

2- What change would make your hometown more appealing to people your age? Use specific reasons and examples to support your opinion.

3- Some people believe that a college or university education should be available to all students. Others believe that higher education should be available only to good students. Discuss these views. Which view do you agree with? Explain why.

4- You have the opportunity to visit a foreign country for two weeks. Which country would you like to visit? Use specific reasons and details to explain your choice.

ENGL 100 Developmental
Exemplification Outline

I. Introduction Style: _____

_____

II. Thesis: (with 3 reasons in parallel structure) _____

_____

_____

_____

_____

III. Topic Sentence #1: (One example of THESIS) is _____

_____

_____

_____

_____

_____

A. What (description): _____

_____

_____

_____

B. Why (connection to THESIS): _____

_____

_____

_____

_____

IV. Topic Sentence #2: (A better example of THESIS) is _____

_____

A. What (description): _____

B. Why (connection to THESIS): _____

V. Topic Sentence #3: (The best example of THESIS) is _____

A. What (description): _____

B. Why (connection to THESIS): _____

VI. Conclusion: (rephrase THESIS):

Appendix 10: Writing Assessment Rubric

	A/P+	B/P	C/P	D/F	F
Elements of the Essay (25 pts) • Introduction uses one of the six recommended techniques. • Thesis is written in the appropriate format (Answer to essay prompt because reason #1, reason #2, and reason #3). • Topic sentences appear at the beginning of the body paragraphs. • Supporting detail/examples/evidence is found within the body paragraphs. • Thesis is restated (using different words from introduction) at the beginning of the conclusion; the meaning remains exactly the same. • Conclusion encourages insight, reflection or action.	**21-25**	**19-20**	**18**	**16-17**	**0-15**
Content/Support (25 pts) • Topic sentences are strong reasons the thesis is correct. • Supporting detail/examples/evidence are specific and relevant. • Supporting detail/examples/evidence are convincing.	**21-25**	**19-20**	**18**	**16-17**	**0-15**
Unity and Coherence (20 pts) • Introduction progresses logically and smoothly from hook to thesis. • Every topic sentence includes a transition word or phrase. • Each detail/example/piece of evidence is introduced with a transition word or phrase. • The body paragraphs are unified. (Discuss one point/reason per paragraph and stick to it. Include only sentences that support the point.) • The restatement of thesis includes a transition word or phrase. • Conclusion progresses logically and smoothly from restatement of thesis to insight/reflection/ charge to action.	**18 - 20**	**16-17**	**15**	**13-14**	**0-12**
Mechanics and Grammar (20 pts) • Errors in punctuation, grammar, spelling, et cetera do NOT interfere with overall communication. • Variety of sentence structures. • Vocabulary, word formations, phrasing do NOT interfere with overall communication.	**18 - 20**	**16-17**	**15**	**13-14**	**0-12**
Quality of Revision (10 pts) • Completes and submits rough draft(s). • Thoughtfully and appropriately responds to instructor feedback from rough draft(s). • Seeks out instructor with any question about his/her comments on rough draft(s). • Willing to make major changes to rough draft(s).	**9 - 10**	**7-8**	**5-6**	**3-4**	**0-2**
Formatting • Prewriting, outline, first draft and any other required supporting material is submitted with final draft in a plastic folder. • Abides by format guidelines in syllabus and/or handout.	**0**	**-2**	**-4**	**-7**	**-10**
Total: /100					

APPENDIX 11: Possible Influence of Outliers on the Correlations

I created boxplots of the change in score by group for both the test and essay scores to check for extreme values. For the test scores, one individual assigned to the experiment group had a change in test score of -3 (shown by a star in Figure A1). In a similar manner, two individuals had a change in essay scores of -5 and 14, respectively, in the experiment group (see Figure A2). In both cases, these extreme observations could have greatly contributed to the overall estimate of the correlation between the test and essay scores in the experiment group. I thus estimated the correlation coefficient again after excluding these three individuals as a sensitivity analysis focusing on the experiment group only.

Figure A1

Boxplot of Change in Score (Test) by Group

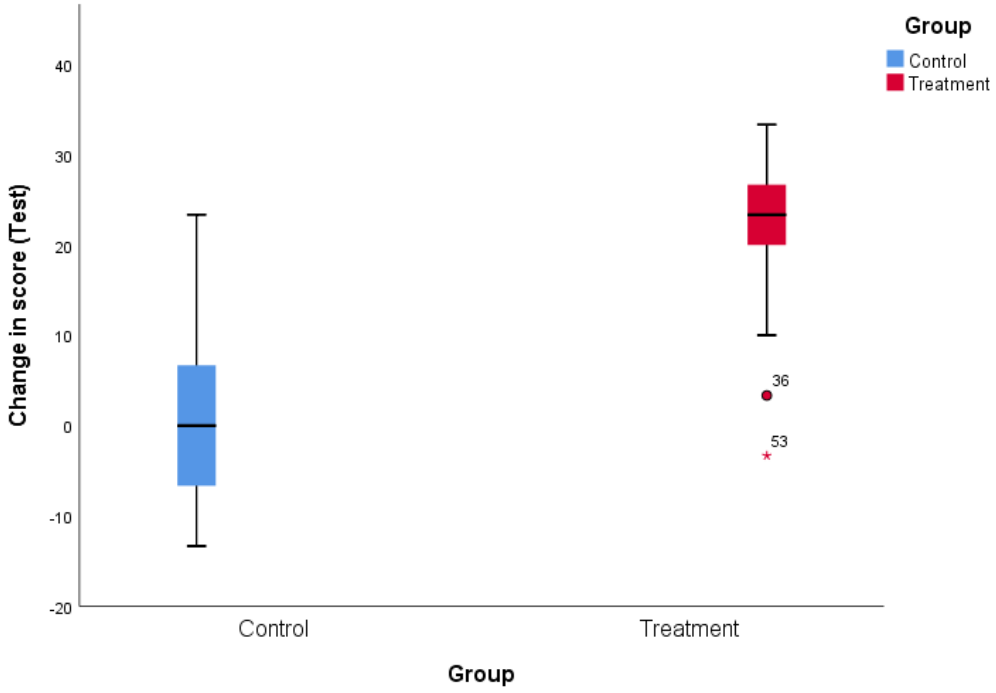
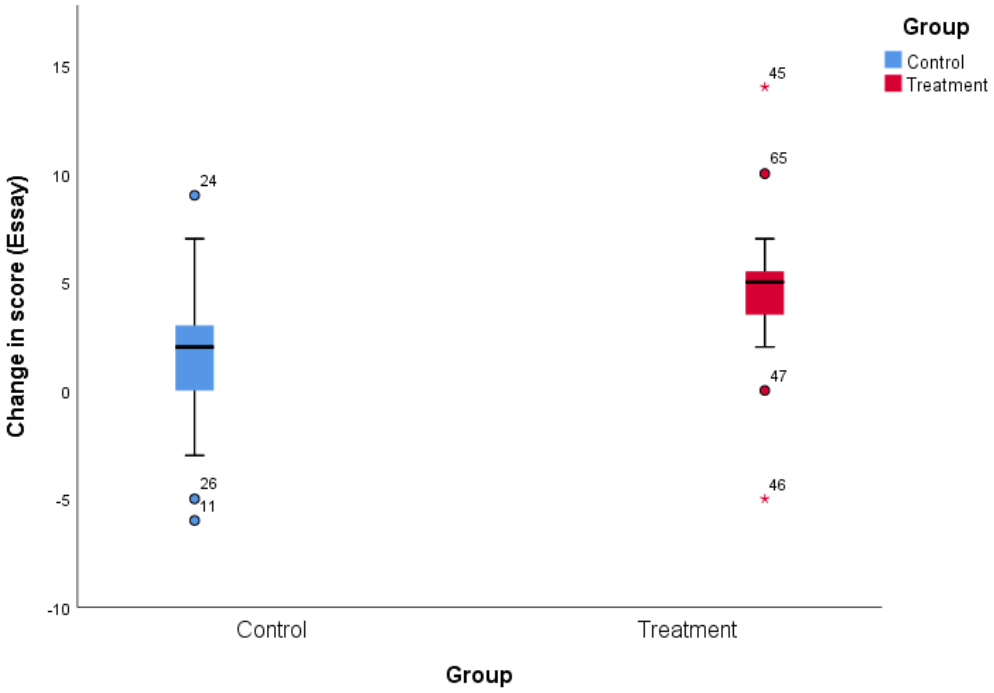


Figure A2

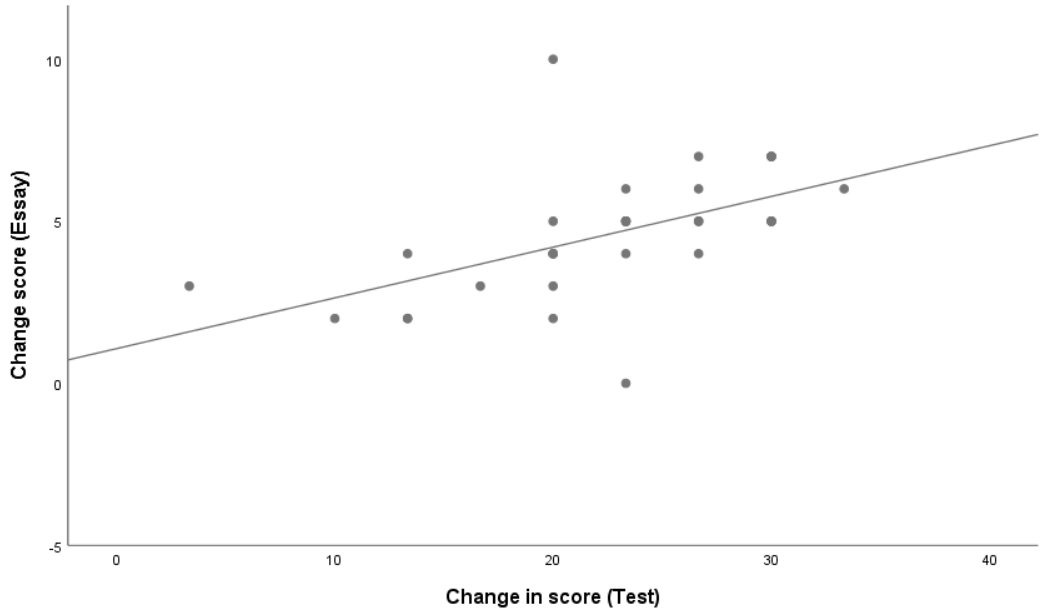
Boxplot of Change in Score (Essay) by Group



There was a stronger positive relationship between the test and essay scores in the treatment group after removing the influential values (see Figure A3). The estimated Spearman correlation coefficient was 0.69 (95% C: 0.48, 0.91, $p<0.001$). I again rejected the null hypothesis that the test and essay scores were not linearly related for the alternative that there was a linear relationship between the two scores in the positive direction.

Figure A3

Scatterplot of Change in Score (Essay) by Change in Score (Test) in the Treatment Group After Removing Outliers



Appendix 12: Participant Information sheet

Participant Information Sheet

Title of Project: The Correlation between Teaching Collocations & Lexical Bundles and Improvement in the Writing skill of First-year University Students

Researcher name: Sally Kondos

Invitation and brief summary:
The proposed research proposal is my final study of my doctoral degree. I have been teaching English as a Foreign Language for 15 years. Over the years, I have noticed that learning a new language is probably a very challenging task . The mechanism of vocabulary learning is still of mystery, but most of the research confirmed that the words are not instantaneously acquired; on the contrary, they are acquired over a long period and from different exposures.
I am proposing a study to investigate how the vocabulary knowledge help with the writing development of First year college student. The programme will be over a period of one semester, which is 15 weeks.
Please take time to consider the information carefully and discuss with your friends if you wish, or to ask the researcher questions.

Purpose of the research:
According to the rules of the Ministry of Higher Education and Research in the United Arab Emirates, all students who desire to learn in any university must pass a standardised test to start their undergraduate studies. If the students are unsuccessful in acquiring the required grade, whether on the TOEFL, the IELTS, or the university Accuplacer exam, they are enrolled in entry level English course which is called ENG 100. It is non- credit course which means students only pass without earning any credits towards their graduation. In the teaching context where the study will take place, the majority of students struggle to score a 6.5 in the writing part of the IELTS exam, or 5 in the TOEFL test. The students' main weakness lies in their inability to elaborate their ideas and develop their writing further; they struggle to expand on a sub-topic in the writing exam. The aim of my study to investigate how vocabulary help the development of the students writing skills.

Why have I been approached?
You have been approached because you are an ENG 100 student and the study participants are 50 firs-year students at an American University in Dubai who are enrolled in English 100 non-credit courses. The sample is representative of the ENG 100 Course population. The English 100 students are fresh graduates of high school with an age that ranges between 18-20 years old. The ENG 100 students spend 4 hours of classes weekly to pick up their language level to cope with the language level of the ENG 101 course; the first credit English course is a series of classes that all students have to take in their freshman and sophomore years.

What would taking part involve?
The study aims to investigate the impact of learning vocabulary on improving the writing proficiency of ENG 100 students. The study will be over 15 weeks. The study participants will meet for an hour four times every week. The participants will write three writing assignments over the study period. The writing assignments aim to analyze the student's ability to demonstrate their knowledge of the

collocations and lexical bundles productively. As mentioned earlier, the study population is freshman students who need to develop their writing further to function well in university courses. The study aims to help the students develop writing skills by teaching them the collocations & lexical bundles. During the intervention study, the student learns the different writing genres, for example, the argumentative genre. Hence, the writing assignments aim to measure if the student understood the collocations & lexical bundles that can be integrated and correctly used in an essay.

I will personally record all the results of the writing assignments and the results of the questionnaires. All the participants will be asked to suggest a pseudonym for themselves that I could use in writing up the forms. I store all the data on the University of Exeter one drive account with a secured password and data encryption. I confirm that I will avoid storing any data on my laptop or USB sticks. I will make sure that the data is not accessible to any third party. I am the only person who has access to it. As soon as I finish my study, I will destroy all the data.

What are the possible benefits of taking part? Research does deliver more comprehensive benefits to ENG100 society, and some indirect benefits might be foreseeable for participants in the study.

What are the possible disadvantages and risks of taking part?
I confirm that there is no personal data involved in my study. The data will not contain any personal information or demographics. The confidentiality of the participants' identities will be assured throughout the research.
Every participant is considered pseudo-anonymized and still identifiable. for example, the essays, as well as the questionnaires, will only contain a number instead of any name. Nontracebaility will be assured throughout the study, and this will be extended to the aggregating of the data.

All the participants' identities will be pseudo-anonymized and still identifiable. I will guarantee that the research will not harm them. I will also guarantee non-traceability in the research. I anticipate no harm in undertaking the study.

Every participant will be assigned a pseudonym throughout the research. The data will be transferred to coded unnamed datasheets.

What will happen if I don't want to carry on with the study?
I will make sure that there is no pressure on the students to participate. Once the participants volunteer, they will be given an ethical consent form explaining all the stages of the programme. They will be assured that they can withdraw from the study at any stage.
The participants can inform of their wish of withdrawing from the study without having to provide any further explanations.
There is no personal data involved in my study. I will delete all the writing assignments from the University of Exeter one drive account.

How will my information be kept confidential?
The University of Exeter processes personal data for the purposes of carrying out research in the public interest. The University will endeavour to be transparent about its processing of your personal data and this information sheet should provide a clear explanation of this. If you do have any queries about the University's processing of your personal data that cannot be resolved by the research team, further information may be obtained from the University's Data Protection Officer by emailing dataprotection@exeter.ac.uk or at www.exeter.ac.uk/dataprotection
The study will take 15 weeks. Once the data is collected and analyzed and my analysis is finalized and approved, I will immediately get rid of the data. I anticipate that the process might take a year long. During that time, the data is stored on the University of Exeter one drive account with a secured password and data encryption. I confirm that I will avoid storing any data on my laptop or USB sticks. I will make sure that the data is not accessible to any third party. I am the only person who has access to it. As soon as I finish my study, I will destroy all the data.
The data will not contain any personal information or demographics.

As a result of this, I confirm that I don't have any particular financial interests in undertaking this research. The research is self-funded, and it is undertaken to complete a degree with no commercial interests
All the participants' identities will be pseudo-anonymized and still identifiable. I will guarantee that the research will not harm them. I will also guarantee non-traceability in the research. I anticipate no harm in undertaking the study.

Every participant will be assigned a pseudonym throughout the research. The data will be transferred to coded unnamed datasheets.

The data will not be used in any other project. I will destroy the data as soon as I finish writing up my findings and finishing my doctoral degree.

Will I receive any payment for taking part?

The research is self-funded, and it is undertaken to complete a degree with no commercial interests

What will happen to the samples I give?

Once the data is collected and analyzed and my analysis is finalized and approved, I will immediately get rid of the data. I anticipate that the process might take a year long. During that time, the data is stored on the University of Exeter one drive account with a secured password and data encryption. I confirm that I will avoid storing any data on my laptop or USB sticks. I will make sure that the data is not accessible to any third party. I am the only person who has access to it. As soon as I finish my study, I will destroy all the data.
I also confirm that the data will not contain any personal information or demographics.

What will happen to the results of this study?

The study is to fulfill my EdD in TESOL . The data will not be used in any other project. I will destroy the data as soon as I finish writing my findings and finishing my doctoral degree.

Who is organising and funding this study?
The research is self-funded, and it is undertaken to complete a degree with no commercial interests

Who has reviewed this study?
This project has been reviewed by the Research Ethics Committee at the University of Exeter and my supervisor Dr Esamaeel Abdollzadeh.

Further information and contact details
Please feel free to contact Dr Esmaeel Abdollzadeh at any time during the project on email e.abdollazadeh@exeter.ac.uk *You can also contact the* College of Social Sciences and International Studies Research Ethics Committee:
ssis-ethics@exeter.ac.uk

Thank you for your interest in this project

Appendix 13: Consent Form

Participant Identification Number: 660005394

CONSENT FORM

Title of Project: The Correlation between Teaching Collocations & Lexical Bundles and Improvement in the Writing skill of First-year University Students

Name of Researcher: Sally Kondos

1. I confirm that I have read the information sheet dated.....05/02/2022..... (version no..1.....) for the above project. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without my legal rights being affected. ☐

3. I understand that relevant sections of the data collected during the study, may be looked at by members of the research team, individuals from the University of Exeter, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐

   I understand that taking part involves anonymised questionnaire responses to be used for the purposes of investigating the impact pf learning vocabulary on improving the writing performance of First Year students ☐
   (e.g.): [inclusion in an archive for a period of up to one year]

   [shared with other researchers for use in future research projects] ☐

   [reports published in an academic publication… ☐

4. I agree to take part in the above project. ☐

Name of Participant	Date	Signature
Name of researcher taking consent	Date	Signature

When completed: 1 copy for participant; 1 copy for researcher/project file

Appendix 14: The Study Dataset

ID	Group	Pre-test essay	Post-test essay	Pre-test MCQ & fill in blanks test	Post MCQ& fill in blanks test	Group	Pre-test MCQ & fill in blanks test	Post-test MCQ & fill in blanks Test	Group	Change score MCQ (test)	Change score (Essay)	
1	Experimental	72	76	13	19	1	43	63	Experimental	20	4	
2	Experimental	73	77	11	18	1	37	60	Experimental	23	4	
3	Experimental	77	82	17	24	1	57	80	Experimental	23	5	
4	Experimental	78	85	10	18	1	33	60	Experimental	27	7	
5	Experimental	79	85	15	22	1	50	73	Experimental	23	6	
6	Experimental	70	73	13	14	1	43	47	Experimental	3	3	
7	Experimental	80	87	17	26	1	57	87	Experimental	30	7	
8	Experimental	78	84	15	23	1	50	77	Experimental	27	6	
9	Experimental	70	73	11	16	1	37	53	Experimental	17	3	
10	Experimental	80	85	15	24	1	50	80	Experimental	30	5	
11	Experimental	73	75	9	13	1	30	43	Experimental	13	2	
12	Experimental	76	80	12	16	1	40	53	Experimental	13	4	
13	Experimental	75	78	11	17	1	37	57	Experimental	20	3	
14	Experimental	74	76	11	15	1	37	50	Experimental	13	2	
15	Experimental	73	87	13	20	1	43	67	Experimental	23	14	
16	Experimental	80	75	15	24	1	50	80	Experimental	30	-5	
17	Experimental	72	72	10	17	1	33	57	Experimental	23	0	
18	Experimental	73	75	12	15	1	40	50	Experimental	10	2	
19	Experimental	80	85	13	22	1	43	73	Experimental	30	5	
20	Experimental	80	85	17	25	1	57	83	Experimental	27	5	
21	Experimental	75	80	13	19	1	43	63	Experimental	20	5	
22	Experimental	79	86	14	23	1	47	77	Experimental	30	7	
23	Experimental	77	82	14	13	1	47	43	Experimental	-3	5	
24	Experimental	76	81	13	20	1	43	67	Experimental	23	5	
25	Experimental	77	82	15	22	1	50	73	Experimental	23	5	
26	Experimental	79	84	16	25	1	53	83	Experimental	30	5	
27	Experimental	80	85	15	22	1	50	73	Experimental	23	5	
28	Experimental	73	77	11	19	1	37	63	Experimental	27	4	
29	Experimental	79	86	17	26	1	57	87	Experimental	30	7	
30	Experimental	77	82	14	22	1	47	73	Experimental	27	5	

31	Experimental	77	81	12	18	1	40	60	Experimental	20	4
32	Experimental	78	82	14	20	1	47	67	Experimental	20	4
33	Experimental	77	83	15	25	1	50	83	Experimental	33	6
34	Experimental	73	75	11	17	1	37	57	Experimental	20	2
35	Experimental	75	85	15	21	1	50	70	Experimental	20	10
36	Control	75	78	15	13	2	50	43	Control	-7	3
37	Control	66	66	10	9	2	33	30	Control	-3	0
38	Control	66	70	11	8	2	37	27	Control	-10	4
39	Control	77	79	12	13	2	40	43	Control	3	2
40	Control	77	74	11	9	2	37	30	Control	-7	-3
41	Control	75	80	14	11	2	47	37	Control	-10	5
42	Control	77	76	12	9	2	40	30	Control	-10	-1
43	Control	80	80	11	9	2	37	30	Control	-7	0
44	Control	80	82	13	17	2	43	57	Control	13	2
45	Control	75	82	14	13	2	47	43	Control	-3	7
46	Control	83	77	8	12	2	27	40	Control	13	-6
47	Control	78	83	15	15	2	50	50	Control	0	5
48	Control	78	80	12	18	2	40	60	Control	20	2
49	Control	77	77	13	14	2	43	47	Control	3	0
50	Control	80	82	14	12	2	47	40	Control	-7	2
51	Control	77	80	12	14	2	40	47	Control	7	3
52	Control	83	80	11	13	2	37	43	Control	7	-3
53	Control	77	79	13	11	2	43	37	Control	-7	2
54	Control	77	82	13	15	2	43	50	Control	7	5
55	Control	80	78	14	13	2	47	43	Control	-3	-2
56	Control	78	79	11	12	2	37	40	Control	3	1
57	Control	77	80	15	13	2	50	43	Control	-7	3
58	Control	78	80	9	5	2	30	17	Control	-13	2
59	Control	70	79	7	9	2	23	30	Control	7	9
60	Control	80	80	12	11	2	40	37	Control	-3	0
61	Control	78	73	16	18	2	53	60	Control	7	-5
62	Control	73	70	13	13	2	43	43	Control	0	-3
63	Control	77	77	14	15	2	47	50	Control	3	0
64	Control	76	80	6	9	2	20	30	Control	10	4
65	Control	73	75	10	17	2	33	57	Control	23	2

Appendix 15: Ethical Approval

Ethics Application: The Correlation Between Teaching Collocations & Lexical Bundles And Improvement In The Writing Performance Of First-year Students In An American University In The Middle East

☐ Approved

Summary | Details | Scope | Methodology | Human Participants | Data Management | Documents | Reviews | **Response** | Versions | Comments

Response

Chair's Action?	Yes
Approval End Date	10 Jan 2023, 00:00
Response	Dear Sally This study is now approved and you are free to commence research. Please be aware that any significant changes to the study should be reviewed by proposing an amendment for review and receiving a favourable opinion prior to implementation. If during the course of the research process you encounter issues or events that significantly change the level of anticipated risks of the research, you should contact the Research Committee for advice. Please remain aware of any UK government guidance or advice if appropriate. Feel free to get in touch if you have any queries. Best wishes for a successful study. Regards Mark Slater Research Ethics Officer

Hello SALLY KONDOS

Welcome to Worktribe.

Recent | Ethics Applications

Latest Announcements

Title	Date
RIME Worktribe Grants Module is now live	31 Jan 2023
Essential Worktribe User Information – Please read	15 Nov 2021

Recent

Ethics Application: The Correlation between Teaching Collocations & Lexical Bundles and Improvement in the Writing Performance of First-year Students in an American University in the Middle East

Risk	Medium
Applicant	SALLY KONDOS
Org Unit	School of Education
Supervisor	Esmaeel Abdollahzadeh
Status	Approved

Ethics Document: Consent Form

Ethics Document: Participant Information Sheet