TURKISH REPUBLIC
TRAKYA UNIVERSITY
INSTITUTE OF SOCIAL SCIENCES
FOREIGN LANGUAGES TEACHING DEPARTMENT
DIVISION OF ENGLISH LANGUAGE TEACHING
A MASTER'S THESIS

THE USE OF PREPOSITIONAL PHRASE BUNDLES IN TURKISH STUDENT ESSAYS: A REGULAR EXPRESSION ANALYSIS

ÖMER FARUK KAYA

ADVISOR
ASSOC. PROF. DR. KUTAY UZUN

EDİRNE 2023

ACKNOWLEDGEMENT

This master's thesis has relied on the culmination of work and effort by an outstanding group of people.

Above all, I would like to express my deepest appreciation to my mentor and advisor, Assoc. Prof. Dr. Kutay UZUN, who not only introduced me to corpus linguistics, but also generously provided a wealth of knowledge on inferential statistics. I would not have undertaken this journey without his mentorship and constructive feedback. I consider myself fortunate to be his student, as he has always been supportive.

It is difficult to overstate my gratitude to Lect. Dr. Hakan Cangır, for his scholarly support, continuing encouragement, and patience. It would have been impossible for me to complete this thesis without his stimulating ideas and insightful guidance. He gave me many hours of his time to give feedback on my academic publications and provided valuable resources on corpus linguistics, helping me to understand computational linguistics and various corpus tools.

I also want to express my appreciation to TUBITAK ARDEB (220K289) for financially funding my studies. In addition, being involved in a corpus-related project enabled me to have wonderful opportunities to attend talks and meet people from other departments and universities. It helped me to fully focus on my research and have access to a wealth of information.

I would like to thank my committee members, Prof. Dr. Muhlise Coşgun Ögeyik and Assoc. Prof. Dr. Taner Can, for the time they have dedicated to reading my thesis and all the valuable comments they have made.

Finally, my sincere thanks also go to Lect. Alper Arslan, for his encouragement, guidance, and book recommendations.

Tez Başlığı: Türk Öğrenci Yazılarında Edat Öbeklerinin Kullanımı: Bir Düzenli İfade Analizi

Hazırlayan: Ömer Faruk Kaya

ÖZET

Geçmiş çalışmalarda anadili İngilizce olan öğrencilerin yabancı dili ingilizce olan öğrencilere kıyasla sözcük öbeklerini daha sık ve daha çeşitli şekillerde kullandığı belirtildi. Fakat, yakın zamanda öğrencilerin yapısal olarak hatalı sözcük öbeği yazım girişimlerinin çekirdek analizi (core expression) kullanılarak tespiti ve bu bulunan sonuçların toplam sözcük öbeği listesine eklenmesi frekans analizlerinin bulgularını daha güvenilir yapacağı savunulmuştur. Türk literatüründe sözcük öbekleri, özellikle edat grubu sözcük öbekleri bu şekilde bir detaylı analizle incelenmemiştir. Literatürdeki bu boşluk üzerine, bu tez Türk üniversite öğrencilerinin yazdığı skorlaması yapılmış fikir denemelerinden oluşan bir öğrenci derlemi (n = 687) kullanarak dört kelimelik edat bazlı sözcük öbeklerini inceledi. Çekirdek analizine (core expression) ek olarak içerisinde iki edat barındıran edat öbekleri düzenli ifade (regular expression) kodları ile tespit edildi. Bu çalışma (a) anadili İngilizce olan öğrenciler (LOCNESS) ile yabancı dili İngilizce olan Türk öğrencilerin (TELC) sözcük öbeklerini karşılaştırmayı, (b) hatalı sözcük öbeği yazımlarını tespit etmeyi, ve (c) dört kelimelik edat öbeklerinin yazma performansına olan etkisini ölçmeyi amaçlıyor. Sonuçlar anadili İngilizce olan öğrencilere kıyasla yabancı dili İngilizce olan Türk öğrencilerin daha fazla sözcük öbeği kullandığını gösterdi, fakat bu kullanımları daha az çeşitlilik içeriyor. Bir sonraki analiz ise Türk öğrencilerin kimi edat öbeklerini fazlaca kullandıkları kimi öbekleri ise yeterinden az kullandığını ortaya çıkarttı. Düzenli ifade analizine göre Türk öğrenciler sözcük öbeklerini genellikle doğru kullanıyor ve en çok sıklıkla kullandıkları sözcük öbeklerinde hata yapıyor. Son olarak regresyon analizi dört kelimelik edat öbeklerinin düşük puanlı öğrencilerin yazma skorlarını tahmin edebileceğini gösterdi. Bu bulgular deyişbilime, sözcük öbeklerinin eğitimine ve ders kitabı düzenlenmesine yardımcı olabilir.

Anahtar Kelimeler: regular expression, sözcük öbekleri, akademik yazma, edat öbeği

Name of Thesis: The Use of Prepositional Phrase Bundles in Turkish Student Essays: A Regular Expression Analysis

Prepared by: Ömer Faruk Kaya

ABSTRACT

Previous studies showed that L1 English students use a larger quantity and a wider variety of lexical bundles than L2 English non-native students. However, recently it has been argued that accounting for attempted bundles identified by core expression analysis can provide a more reliable overview of total frequency counts. Such analysis on preposition phrase-based bundles remains a research gap in the Turkish context. Considering the research gap, this thesis investigates 4-word preposition phrase-based lexical bundles in a learner corpus of human-rated L2 English opinion essays (n = 687) by Turkish university students. As an extension to the core expression approach, this study uses a set of regular expression codes to identify bundles with two prepositions. The present study aims (a) to compare lexical bundle productions of English native (LOCNESS) and Turkish non-native writers (TELC), (b) to identify erroneous and attempted uses of bundles, and (c) to examine the relationship of 4-word preposition-based lexical bundle use to writing performance. It was found that TELC contained a larger number of bundles, but the use of bundles was less varied than LOCNESS. Statistical analyses revealed a pattern of overuse and underuse of target bundles. Turkish students exhibited a high level of accuracy in the use of bundles and most errors were related to the bundles they most frequently use. Lastly, regression analysis revealed that preposition phrase-based bundle frequency could predict writing scores of low-scored essays. These findings have several implications for future L2 phraseology research, instruction of lexical bundles, and coursebook design.

Keywords: regular expression, lexical bundles, academic writing, preposition phrase

TABLES OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

EFL: English as a Foreign Language

ESL: English as a Second Language

L1: Native Language of the Speaker

L2: Second / Foreign Language of the Speaker

LB: Lexical Bundle

LOCNESS: The Louvain Corpus of Native English Essays

NNS: Non-Native Speakers

NP: Noun Phrase

NS: Native Speakers

PP: Prepositional Phrase

SLA: Second Language Acquisition

TELC: Turkish English Learner Corpus

VP: Verb Phrase

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1. Background to the study

Each year, millions of students enroll at universities, and most universities evaluate and assess students based on their ability to write essays, summaries, and other types of extended texts. For this reason, there is a premium placed on acquiring proficiency in academic writing since it is a key to academic success and the work that follows it (Kellogg & Raulerson, 2007). However, it is a well-known fact that writing is an activity that is challenging for many (Hyland & Tse, 2007; Wood, 2015). There is converging evidence in the literature to suggest that both expert and novice Turkish learners of English have difficulty in writing skills, especially academic writing (e.g., Altınmakas & Bayyurt, 2019; Demirel, 2017; Ertürk & Öztürk, 2022; Köse et al., 2019). Vocabulary has long been a topic of research on writing performance. While early vocabulary studies have focused on individual lexical items, psycholinguistic and corpus-based inquiry have challenged the idea that language is not just strings of individual words; investigated vocabulary also has its preferred lexico-grammatical company. These corpus studies provided empirical evidence that language is not merely a construct of independent single words but also consists of grammatical structures that frequently occur together in systematic statistical patterns. The studies of multi-word strings that have a statistical tendency to occur together were collected under the umbrella of formulaic language studies. Formulaic language encompasses a wide range of structures including idioms, phrasal verbs, and collocations, which are well-recognized examples of these language patterns.

Formulaic language is poorly defined in the literature, creating ambiguity in its definition and identification (Weinert, 1995). According to Weinert, one distinctive characteristic of multi-form strings is that they are stored and produced as whole chunks, similar to the mental processing of single words. To clarify, although these structures are combinations of two or more words, they are processed and encoded in the mind as whole units. Despite the lack of consensus on its definition, formulaic language has been

extensively analyzed, especially in the domain of corpus linguistics. By using a corpus (i.e., a large representative database of spoken or written texts) researchers have paved the way for more general hypotheses about language by analyzing databases comprised of millions of words (e.g., Biber et al., 1999). It was discovered that formulaic language is highly prevalent in both spoken and written language. To provide evidence, Erman and Warren (2000) calculated that formulaic language accounts for one-third to one-half of it. In a similar study, Biber et al. (1999) found that one-fifth of the total words in academic prose are formulaic language. Biber et al. also demonstrated that the use of formulaic language has the potential to distinguish between different types of discourse. To illustrate, the bundle *in the case of* and *on the other hand* are typical of academic discourse as they frequently occur in academic writings than in conversations (Biber et al., 1999). The use of formulaic language can serve a variety of functions (Schmitt, 2005) in addition to occurring in predictable contexts. They can be used as discourse organizers (e.g., *on the other hand*) or signal rhetorical moves (e.g., *it is unsure for*).

Given its high frequency of use, researchers began to acknowledge formulaic language as an important building block of discourse over the past thirty years. As a result, formulaic language has become an independent field of study. Though the number of formulaic language studies increased in recent years, the popularity of formulaic language research has led to confusion due to the lack of consistent terminology used by researchers to describe formulaic language. To illustrate, Biber et al. (1999) used the term "lexical bundle", whereas Erman and Warren (2000) took “prefabs” as a working definition for formulaic language. Lexical bundles are recurring multi-word expressions of three or more words that commonly co-occur in a register (Biber et al., 1999). Identifying lexical bundles requires a set of measures for frequency and dispersion criteria for word occurrence. Since lexical bundles can be extracted solely using frequency measures, they are methodologically straightforward. Research on lexical bundles generally involves structural or functional classification of bundles. For example, Ädel & Erman (2012) classified LBs according to their functions (e.g., discourse organizers). Then, based on the frequency of LBs carrying those functions, they compared the LB preferences of native-speakers (NS) and non-native speakers (NNS). The results indicated that non-native speakers used fewer and less varied lexical bundles than native speakers.

Studies focusing on the frequency of lexical bundles have presented noteworthy results, but it should be noted that the appropriate use of lexical bundles is as important as the production of lexical bundle types and tokens. According to Shin et al. (2018) and Huang (2015), the accuracy of lexical bundles in learner texts has not been fully addressed in previous studies. Shin et al. (2018) argued that solely relying on frequency-driven analyses without investigating the use of attempted bundles (i.e., lexical bundles that are correctly used in context, but erroneous in form) can lead to unreliable results on the overuse and underuse of lexical bundles. Although frequency-driven or corpus-driven methods allow for a quick inquiry of the language patterns without requiring much manual work, they do not adequately draw an inclusive interpretation of lexical bundle use as they cannot detect the attempted use of LBs. To address this problem, Shin et al. (2018) coined the term core expression to refer to "a phrase formed around a core word that constitutes an important component, often lexical, of each LB, along with a following word" (p. 32). To give an example of a core expression: *long time* is the core expression of *in a long time* and *for a long time*. When the target lexical bundle list is retrieved, a manual check of "core expressions" ensures that attempted LB uses by learners are counted as well. Despite its assumed benefits, a few studies have used core expression analysis to examine the erroneous uses of function words (e.g., articles) embedded in lexical bundles (e.g., Lee et al., 2020; Shin, 2018; Shin et al., 2018; Uzun, 2018). To the author's knowledge, Lee et al. (2020) is the only study to use core expression analysis to investigate prepositions found in four-word lexical bundles. Therefore, a detailed accuracy analysis of PP-based bundles, and how attempted bundles can contribute to overall bundle frequency types and tokens remains a gap in the literature, especially in the Turkish context.

Prepositions are known to be challenging for second language learners (Gass et al., 2014). In addition, Shin et al. (2018) demonstrated that a considerable proportion of learner errors in lexical bundles are due to incorrect use of prepositions. Also, several studies (e.g., Shin, 2019; Yoon & Choi, 2015) reported that NNS university students use preposition phrase (PP)-based bundles less frequently than native speakers. Since the English learners considered in this thesis are university students, the importance of investigating preposition phrase-based lexical bundles becomes even more pronounced.

It is also pointed out in a line of research that there is a relation between formulaic language use and proficiency (e.g., Appel & Wood, 2016; Chen & Baker, 2010; Huang, 2015; Staples et al., 2013). Considering the literature gap, the main goal of this thesis is to investigate the use of PP-based lexical bundles by Turkish university students. As an extension to Lee et al. (2020), the present study aims to test the function of regular expression analysis for analyzing erroneous use of lexical bundles as regular expression analysis can help spot erroneous uses of prepositions and articles regardless of their position. Therefore, to fully explore the learner errors and the use of PP-based lexical bundles, the present study used a set of regular expression codes.

1.2. Significance of the study

The present study aims to give a comprehensive overview of how Turkish university students use PP-based bundles in their argumentative essays. In this regard, the main objectives of this thesis are to detect the overuse and underuse of lexical bundles, along with the erroneous uses. Identifying the problem area is only the first step towards the solution. Formulaic language studies do more than describe the use of word combinations and retrieve relevant frequency lists; they also provide insights into pedagogical issues of teaching, testing, and acquiring them. The potential contribution of the analysis of PP-based bundles to language pedagogy can be discussed within the domains of applied linguistics and second language acquisition (SLA).

Various teaching activities can benefit from the data on the frequency and type of PP-based lexical bundle errors. For example, Musgrave and Parkinsons (2014) designed a consciousness-raising task to develop students' awareness of the form and meaning of noun-noun phrases. Based on the overuse of prepositions, designing a similar activity can help students use language more accurately. Moreover, one reason for English as a foreign language (EFL) learners' overall lack of success in using lexical bundles accurately might be the scarcity of natural input. Explicit instruction on meaning relationships expressed by PP-based LBs can yield promising results as they are argued to be more effectively taught explicitly (see Schmitt, 2005). Additionally, linguistic structures, such as articles and prepositions, are often taught in language classes without proper context. Lack of

context prevents students from fully understanding how linguistic structures operate, leading to a potentially incomplete understanding of the language feature being taught. Therefore, employing a pre-test and post-test design Shin and Kim (2017) tested the effectiveness of core-expression-based article instruction by providing explicit instruction on adjoining articles embedded in lexical bundles. Shin and Kim observed a gradual decrease in omission errors related to definite article use at the end of the study.

Corpus-based evidence not only directly contributes to language teaching, but also has indirect implications. For instance, corpus data from NNS on preposition errors can affect the selection, description, and sequencing of assessment materials by providing information about items learners misuse, underuse, or overuse. For example, the information provided by the comparative analysis of learners at various levels can help the construction of the test items to be more consistent with the proficiency levels. Also, data on typical errors can assist item analysis as it can help test designers select suitable distractors that can attract students who do not know the correct answer. Statistically weighted information on error distribution can help design tests that can distinguish between low and high-proficiency learners more reliably. The influence of error-annotated corpora is also reflected in the syllabus and coursebook design. For instance, the Italian version of *The English in Mind* series contains "Get it Right!" tip boxes, which provide authentic examples of typical Italian learner errors (Granger, 2015b).

In a similar fashion to the previously mentioned examples, current study has the potential to make significant contributions to the field and literature. For example, the bundles identified to be overused or underused can serve as a useful resource for instruction. In the case of overuse, teachers can design activities to instruct learners on the alternatives to these bundles (see Alhassan, 2018). The findings on the erroneous uses of bundles might encourage practitioners, especially the instructors in universities teaching preparatory year language courses, to put more emphasis on the bundles that learners commonly misuse. The relation of PP-based bundle frequency and writing performance can contribute to the research as most studies on four-word lexical bundles do not utilize regression analysis. Score prediction models on PP-based bundles can be valuable, especially for assessment purposes, since they can help calibrate automated essay-scoring algorithms. As a final note, the present study is one of the few to use regular

expression analysis to analyze the partial production of LBs. Therefore, it can promote more researchers to use regular expression analysis to analyze the use of reference bundles.

1.3. Purpose

The purpose of the research is to find reliable empirical evidence that can shed light on Turkish students' use of PP-based lexical bundles in academic writing. The investigation of erroneous uses of PP-based bundles is a research gap to fill in the Turkish context. Accordingly, the aims of this research are as follows: (1) to compare the use of PP-based lexical bundles between opinion/argumentative essays written by Turkish EFL students and native writers (2) to investigate the erroneous uses of PP-based lexical bundles in Turkish EFL students' essays and (3) provide insights into the relation of PP-based bundles to writing performance. To this end, the following research questions will guide the study:

RQ1: Is there a difference between L1 Turkish L2 English students' use of 4-word PP-based bundles compared to those produced by native speakers of English?

RQ2: How accurately do L1 Turkish L2 English students use PP-based lexical bundles? How do they use the regular expressions of each lexical bundle?

RQ3a: Does 4-word PP-based bundle frequency predict the writing performance of Turkish students?

RQ3b: Does the frequency of 4-word PP-based differ between high-scoring and low-scoring essays?

1.4. Research outline

This thesis is structured in six chapters. Following the introductory chapter, chapter 2 outlines the necessary background information on formulaic language and lexical bundles, along with the relevant literature. Chapter 3 describes the tools for lexical bundle extraction, and the details of the research procedure followed. Chapter 4 presents the results obtained from the investigation of three exploratory research questions, and

Chapter 5 will provide a comprehensive discussion of these findings. Lastly, in Chapter 6, a summary of the findings and their implications for teaching and language research will be presented.

1.5. Definitions

Corpus: Corpus is "a collection of machine-readable authentic texts (including transcripts of spoken data) which is sampled to be representative of a particular language or language variety" (McEnery et al., 2006, p. 5).

Lexical bundles: Lexical bundles are recurring sequences of three or more words that commonly co-occur in a register (Biber et al., 1999).

Regular expression: The regular expression, or regex, is a computational pattern matching mechanism that uses PERL codes to search for and extract specific language patterns.

Token: The frequency of the number of lexical bundles.

Type: The number of unique lexical bundles.

CHAPTER 2: LITERATURE REVIEW

This literature review begins with a definition of corpus and formulaic language to situate the present study within lexical bundle research. The chapter then presents an operational definition of lexical bundles and the relevant previous studies of lexical bundles to provide a rationale for the exploratory research questions posed in the introductory chapter.

2.1. Corpora and corpus methods

In linguistics, corpora and corpus methods have extensively been used since the 1960s (Altenberg, 1991; Granger, 1998), presenting samples of language use and variation in the form of data and context, which are stored and analyzed using computer technologies. In a constituted definition, corpus refers to a large collection of authentic texts stored electronically, selected systematically to represent a particular language use or variety (Biber et al., 1998; McEnery et al., 2006; Stubbs, 2004). More specifically, corpus has the following characteristics:

- large: millions, or even hundreds of millions, of running words, usually sampled from hundreds or thousands of individual texts;
- computer-readable: accessible with software such as concordancers, which can find, list and sort linguistic patterns;
- designed for linguistic analysis: selected according to a sociolinguistic theory of language variation, to provide a sample of specific text-types or a broad and balanced sample of a language (Stubbs, 2004, p. 106).

Using corpora and corpus methods, researchers can conduct analyzes on a larger scale and tackle previously challenging research questions. The main advantage of corpus-based methods is the ability to analyze a much wider range of language data automatically or semiautomatically. With corpus technologies, language use can be examined in greater detail than with other approaches. In a frequency-driven analysis, Coxhead (2000) identified the individual vocabulary units typical of academic writing by compiling a

corpus (3.5 million tokens) of various academic texts (more than 400 texts) to create an academic vocabulary list which can students who are studying at tertiary education in English.

In addition to basic counts of linguistic features, corpus-based analyzes can also go beyond this and examine more complex patterns and relationships within language (Biber et al. 1998). According to Biber et al. there are linguistic associations of language features, which can be divided into two categories: lexical associations and grammatical associations. The statistical likelihood of the word *morning* occurring together with *good* can be given as an example of lexical associations while that- complement clause to complete the meaning of verbs (e.g., *demonstrated that*, *showed that*) is an example of grammatical associations. On the contrary, Chomsky (1956) argues that language has immense creative power, and it is possible to create infinite sets of sentences within the boundaries of grammar rules. However, it is unlikely for a native speaker to produce *great morning* instead of *good morning*. Based on the frequency of use data and the statistical likelihood of words occurring together, it is fair to say that native speakers do not exercise the creative potential of language to its full extent (see Pawley & Syder, 1983). Instead, there is a set of commonly used expressions that are conventionalized and are often used together. Moreover, Erman and Warren (2000) and Biber et al. (1999) presented corpus evidence suggesting that these multi-word units make up a large proportion of language use. The complex relation and interaction between grammar and the lexicon, as well as the processing of multi-word units in mind, are explored in the domain of phraseology or formulaic language studies.

2.2. Formulaic language

Hyland and Tse (2007) have cited some studies (Campion & Elley, 1971; Coxhead, 2000; Nation, 1990; Praninskas, 1972) showing that the investigation of vocabulary necessary for academic studies has long been an important issue. With the widespread availability of corpus and innovative corpus technologies, the 2000s and beyond saw many studies that promoted the idea that language is not just strings of individual words (e.g., Bestgen, 2017; Bestgen & Granger, 2014; Wray, 2002). These

studies suggested that investigated vocabulary also has its preferred lexico-grammatical company and investigation of multi-word structures can be as valuable as studies on single words. In other words, besides being governed by grammatical and semantic rules, language use also largely relies on formulaic language. Formulaic language is a commonly used term for the combination of two or more words that are stored and retrieved from memory as a whole chunk. To elaborate more, Wray defines formulaic language as:

> "A sequence, continuous or discontinuous, of words or other meaning elements, which is, or appears to be, prefabricated: that is stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar." (p. 9).

Wray's definition successfully captures the basic criteria for defining the complex nature of this phenomenon by illustrating that formulaic language is an umbrella term covering a wide range of multi-word units such as collocations, idioms, lexical bundles, phrasal verbs, and p-frames. What contributed to this diversity of terminology is the expanding body of literature over the past few decades and the development of new approaches to identifying formulaic language. Although theories on the tendency of words to occur together can be traced back to the mid-1700s (Schmitt, 2022), formulaic language studies have only recently gained popularity with the development of corpus technologies and recent language and learning theories.

2.3. Historical background of formulaic language

Linguistic studies on cognitive perspective (Becker, 1975; Bybee, 2006; Langacker, 1987, 2000; Tomasello, 2000), emphasis on frequency in language acquisition (Tomasello, 2003) and empirical investigations of the structure of language employing corpus method (Biber et al., 1999; Erman & Warren, 2000; Pawley & Syder, 1983; Sinclair, 1991) has caused an emerging interest in formulaic language studies. However, despite its popularity in recent years, it was not until the 1970s that linguistic inquiry attention to these multi-word combinations (Wood, 2015; Yoon, 2021). The reason for

this might be theoretical assumptions of language at the time and the lack of computer technology to perform empirical analysis on large sums of data.

In the second half of the 20th century, Chomsky's critique of Skinner's *Verbal Behaviour* (Chomsky, 1964) brought cognitive processes to language learning to the forefront and attracted more researchers to investigate the systematic nature of learner language focusing on morpho-syntactic features (Weinert, 1995). Chomskyan view of language (Chomsky, 1956, 1967) rests on two ideas: language is an innate ability, and it is based on a collection of logical rules (i.e., grammar) that allows speakers to generate an infinite number of novel sentences. In more detail, this approach holds that the language is made up of empty slots that can be filled by any word as long as grammatical rules are followed. Words were seen only as formal elements to form sentences, without any interaction with syntax. Chomsky's ideas on the innateness of grammar and grammar being the only restriction on language production were soon rejected by studies on formulaic language. In the same period, studies on formulaic language (e.g., Becker, 1975; Firth, 1957; Pawley & Syder, 1983) argued that morphosyntactic rules are not the sole determinants of how words are combined, and vocabulary and syntax interact in complex ways. The work of Firth (1957), for example, is regarded as a pioneer of the frequency-based approach to language analysis (Wood, 2015). As opposed to the Chomskyan view of language, Firth (1957) claimed that it is possible to determine the meaning of a word by the words surrounding it with his famous statement "you shall know a word by the company it keeps" (p. 11). That is, there is a statistical likelihood of words occurring together. For example, the word *traffic* is more likely to have *heavy* as an adjective than *intensive*. The ideas of Firth (1957) had a lasting impact on the field of linguistics and influenced the way that language is understood, leading to further research on how various word combinations are processed and used. Similarly, Becker (1975) suggested that understanding how phrases function is fundamental to understanding the language since most utterances are generated in predictable social settings and require only the selection of appropriate formulas, clichés and other fixed expressions stored in mind. Becker supported his ideas by providing examples of idioms and common expressions and observations of this kind have inspired an extensive number of studies (e.g., Pawley & Syder, 1983; Sinclair, 1991).

Studies in the following years emphasized the mental processing advantages of formulaic sequences by arguing that learners' fluency is determined not by their knowledge of grammar rules, but by the amount of formulaic language they retain in their memory. In a seminal work, Pawley and Syder (1983) suggested that the average native speaker knows hundreds of thousands of lexicalized sentence stems, which explains their language fluency. To further emphasize the restraints on language creativity beyond grammar rules, Sinclair (1991) proposed two seminal concepts the "open-choice principle" and the "idiom principle" - two factors influencing the choice of words. The open-choice principle resembles the Chomskyan perspective of grammar because it refers to the random selection of individual words to fill the open slots in formulaic sequences with the only restriction being grammar rules. It is also called the "slot-and-filler" model because a number of lexical items can fill the slots found in clauses, phrases and many other units. According to the idiom principle, "a language user has available to him a large number of semi-preconstructed phrases that constitute single choices" (Sinclair, 1991, p. 110). Therefore, in contrast to choosing lexis to fill the slots, the idiom principle suggests the choice of two words at once is possible and just like single words, formulaic language is processed and stored in our minds as whole units. In the idiom principle, Sinclair also argued that words do not occur randomly in a text and there are more restraints beyond grammar; some words have a phraseological tendency to occur together, and the register and context of discourse affect the choice of words. As exemplified in Sinclair, the word *happen* is usually used in a certain semantic environment that connotates unpleasant things such as accidents. Rather than the choice of lexis to fill the slots, the idiom principle suggests that the choice of a simultaneous choice of two words is possible. Technological advancements and corpus use enabled more researchers to uncover empirical evidence supporting the pervasive use of formulaic language in conversation and written discourse (such as Altenberg, 1998, and Biber et al., 1999). The analysis of formulaic language can be done more quickly and precisely when corpus and concordance technologies are used. To give examples of early frequency-driven analyzes, Erman and Warren (2000) calculated that formulaic language covers between one-third to one-half of both spoken and written language. Also, based on *Longman Spoken and Written English Corpus* (40 million words), Biber et al. (1999) found that formulaic

language accounts for 28 per cent of conversation and 20 per cent of academic writing. In addition, recent technologies that capture auditory and visual responses to formulaic language allowed a more in-depth understanding of how the formulaic language is mentally processed (e.g., Conklin & Schmitt, 2008; Siyanova-Chanturia et al., 2011; Sosa & MacFarlane, 2002).

Underlying these ideas and findings about formulaic language usage-based perspective provides a theoretical ground for the processing and use of formulaic language. Usage-based perspectives on language explore how frequency and repetition affect, and ultimately bring about, form in language, and how this knowledge affects language comprehension and production (Ellis et al., 2008). According to usage-based accounts of language, people possess linguistic skills in the form of a "structured inventory of symbolic units" as a result of their cumulative experience with language across their lifetime (Langacker, 1987). In other words, language acquisition is a cognitive organization of language usage and experiences. The accumulation of usage-based language studies, most of which presents corpus-based evidence, has produced evidence that language processing is highly affected by formulaicity, and that language use substantially relies on formulaic language (see Ellis et al., 2008 for an overview). Further, Ellis (2014) argues that one of the key principles and objectives of instructed second language learning is building a rich inventory of formulaic language and rule-based competence since this repertoire of formulaic language skills can satisfy the fluency and functional needs of the language. All in all, Formulaic language is now relatively well established as one of the most fundamental components of language by scholars that take a usage-based emergentist perspective on language since it is pervasive in language use, is a basic unit of meaning and functions, and has mental processing advantages compared to non-formulaic language (Cangır, 2018).

2.3.1. Approaches to formulaic language analysis

There has recently been a significant amount of research on formulaic language, particularly corpus-based research. In light of these developments, various methods to identify formulaic language were developed, such as phraseological and frequency-based approaches. According to phraseological approaches, formulaicity can be defined in two ways: first, as the extent to which a word combination can be inferred from its parts, and

second, as the interchangeability of words with similar meanings (Durrant & Mathews-Aydinli, 2011). Durrant and Mathews-Aydinli claims that phraseological approaches are mainly concerned with the non-compositionality of formulaic language. Non-compositional formulaic language usually has pragmatic meaning and does not allow for decoding. Idioms and sayings (e.g., *all that glitters is not gold*) can be counted as examples of non-compositional formulaic language because they do not allow for syntactic manipulations as it changes the meaning and function.

Frequency-based approaches, on the other hand, rely on the statistical identification of formulaic language found in corpora (Wood, 2015). Statistical measures and automatic methods of language analysis such as n-gram analyzers and concordancers are foundations of frequency-based approaches. Since the criteria and methods applied to identifying formulaic language vary, the researchers focus on different types of word combinations. According to Wray (2002, p. 9), there are more than forty terms to describe the different aspects of formulaicity, and the list includes n-grams, p-frames, chunks, collocations, and idioms. These terms are distinct from one another in terms of length, fixedness, compositionality, and institutionalization (see Schmitt, 2005). An idiom, for example, is generally fixed in form and meaning (e.g., *kick the bucket*), while an n-gram is more variable in verbs, prepositions and content words (e.g., *is because of*). Also, n-grams are far more common than idioms.

2.3.2. Formulaic language studies

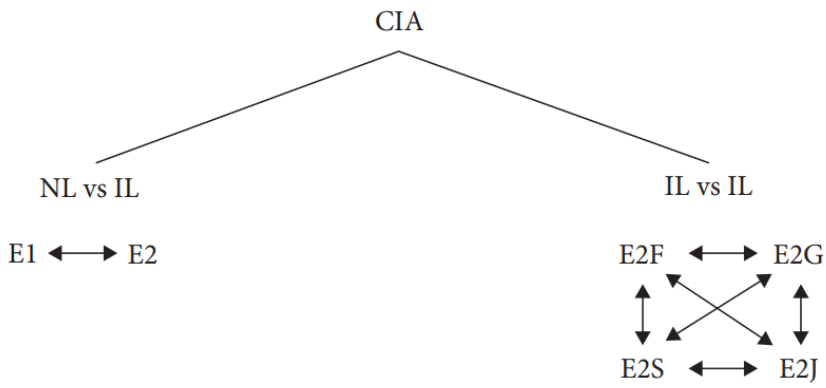
The studies on academic writing mentioned above fall within the domain of Learner Corpus Research, as they utilize a compilation of texts gathered from language learning (i.e., learner corpus). Learner Corpus Research (LCR) is a subfield of corpus linguistics that emerged in the 1980s and focuses on texts gathered from language learners (i.e., learner corpus). The International Corpus of Learner English (ICLE; Granger, 1993; Granger et al., 2020) can be counted as one of the earliest examples of learner corpora. ICLE is meticulous work of scholarship, which encompasses information about the learners (e.g., L1 background and age) as well as task-related details (e.g., topic, use of reference materials). This tradition of rich metadata has been followed by many learner corpora created after ICLE. Furthermore, Sinclair (1991, p. 9) asserts that “the results are

only good as the corpus". It is fair to say that better control of confounding variables learner corpora leads to greater generalizability of findings and more reliable analysis.

Thanks to the widespread availability of corpora like ICLE, it is possible to analyze learner corpora using an analytical procedure called Contrastive Interlanguage Analysis (CIA). Learner corpus and CIA allow researchers to survey the differences between learners' interlanguage and native speaker performance as well as the effect of intergroup differences among learners (e.g., proficiency, L1 background, academic experience).

Figure 1

Contrastive Interlanguage Analysis Model by Granger (1996)



According to Bestgen and Granger (2014), CIA is concerned with comparing learner corpus to native counterparts to uncover distinguishing features of learner language or with other samples of learner data to assess the generalizability of findings. Figure 1 illustrates two types of comparisons for CIA:

- Type 1: comparisons of native and learner varieties of one language, or
- Type 2: comparisons of a target language such as English produced by learners from different L1 backgrounds (Granger, 1996, p. 43).

In the figure "IL" refers to interlanguage and "NL" stands for native language or the target language. Therefore, the first type is NS and NNS comparisons, which can identify distinguishing features of NNS learner writing. This type of comparison helps

identify the variations (i.e., overuse and underuse) in the use and frequency of words, formulaic language, and other structures. In Ishikawa (2011), a comparison between the English argumentative essays written by Japanese learners of English and that of English native speakers revealed differences in the use of n-grams between NNS and NS. The study investigated the use of 2–, 3-, and 4-word n-gram and found that the NNS overused first person pronouns, such as subjective interpersonal metaphors (e.g., *I think that, I don't think*) while they underused preposition phrases. The results indicated that NNS use of n-grams is not typical of argumentative writing because instead of discussing a point their arguments were centered around themselves and their opinions. Prepositions are considered as important constructs for presenting arguments and increasing information density (see Biber et al., 2004), and Ishikawa revealed that underuse of such features indicates a potential area of difficulty faced by Japanese learners. Additionally, Geluso (2022) conducted an exploratory study comparing argumentative essays in a Spanish corpus of non-native learners and the L1 English corpus to investigate the use of prepositions in p-frames, a type of recurring formulaic language that contains slots (e.g., preposition + the * of). The study revealed that there was a significant difference in rates of recurrence of p-frames, especially the frame "*on the * of*". In the English language, "*on*" in phrases such as "*on the plane*" conveys the idea of being inside a plane, but the Spanish word "*en*" which directly translates to English "*in*" is used for communicating the same idea (Geluso, 2019). Therefore, the study provided important insights into the L1 influence. As can be seen, such comparisons employing contrastive interlanguage analysis can provide details on the use of language features.

The second type, comparisons across non-native speakers, aim to distinguish those features of L2 English that were L1-dependent (Granger, 2015a). Granger argues that L2-L2 comparisons should be interpreted beyond L1 background to include variables such as proficiency levels, educational setting, and task effects. A learner corpus representing learners at different proficiency levels can be a valuable resource for researchers interested in investigating the development of L2 proficiency. A study compared the use of n-grams across different proficiency levels in a corpus of Asian EFL learners, which is scored based on high-stake proficiency tests (Chen, 2019). Chen sought to test the predictive power of formulaicity by analyzing the relation of delta p mean

scores to n-gram use and learner proficiency. It was found that learners at higher levels are able to produce more complex phrases and expressions and there is an increase in formulaicity with proficiency. The study indicated that there is a relation between formulaicity and proficiency, highlighting the importance of phraseological competence in assessing language proficiency.

In a contrastive interlanguage analysis using a learner corpus, it is possible to examine both overt and covert differences between interlanguage performance and native speaker performance. Examining the characteristics of the learner language can have potential implications for language teaching and provide new insights into how languages are learned (Hasselgård & Johansson, 2011). On these premises, a significant body of research has been produced in the field of LCR using CIA methodology to analyze learner language. Therefore, the next section is dedicated to discussing the L1-L2 and L2-L2 comparisons of lexical bundle use in terms of the total frequency (tokens) and varied use (types) of lexical bundles and the structural and functional variations.

2.4. Lexical bundles

First coined by Biber et al. (1999), lexical bundles are recurring sequences of three or more words that commonly co-occur in a register. As Cortes (2004, p. 400) notes, “Many lexical bundles are not idiomatic: rather, their meaning is transparent, fully retrievable from the meaning of the individual words that make up the bundle”. Lexical bundles are identified based on a frequency-based approach (Biber & Barbieri, 2007), which often solely relies on statistical measures and frequency counts. In fact, Hyland (2012) asserts that "bundles are statistically the most frequent recurring sequences of words in any collection of texts" (p. 150). This claim is supported by Biber et al. (1999), which report that lexical bundles constitute 28 per cent of conversation and 20 per cent of academic prose in *Longman Spoken and Written English* corpus (around 40 million words corpus). Thus, recurrence is one of the defining characteristics of lexical bundles, and to be counted as recurrent, lexical bundles should satisfy frequency and dispersion criteria.

The frequency threshold criterion determines the number of lexical bundles included in the analysis and varies depending on the target lexical bundle length and corpus size. The normalized frequency threshold for large corpora often ranges between 10-40 per million words (e.g., Biber et al., 1999; Biber et al., 2004), while for relatively small spoken corpora the frequency threshold is usually lower than ten times per million words (e.g., Lee et al. 2020). To ensure that the identified lexical bundles are representative of the corpus, they must appear in multiple texts. Setting a dispersion threshold helps to guard against the influence of a single writer's idiosyncratic style, thereby providing a more accurate depiction of the language used across the corpus. For example, Biber et al. (1999) only considered multiword units appearing across at least five texts as lexical bundles.

Lexical bundles are often incomplete units that are usually three-word, four-word or five-word structures. For example, the four-word lexical bundle *in terms of the* is a combination of a noun phrase and a prepositional phrase, in which the prepositions modify the noun *terms*. Though only 5% of academic LBs are complete structures, they can serve as discourse signaling devices (Biber et al., 1999). For example, *similar to those of* is used for comparing examples or events. There are many examples of lexical bundles that can be found in academic contexts (e.g., *as a result of, in a number of, on the other hand*) and conversation (e.g., *you know I mean, she said to me*) that mark discourse. Since LBs signal discourse and are highly prevalent in language, researchers attempted to classify and label structural types and discourse functions of these lexico-grammatical units.

2.4.1. Structural and functional taxonomies

In the study of lexical bundles, there is a growing amount of attention on the analysis of discourse functions and structural characteristics of lexical bundles (e.g., Biber et al. 1999; Biber et al. 2004; Biber & Barbieri, 2007; Cortes, 2004; Hyland, 2008; Simpson-Vlach & Ellis, 2010). Biber et al. (1999) described the use of lexical bundles in both spoken and written registers, drawing on a large corpus that covered a wide range of topics and genres (e.g., academic journal articles, newspaper reports, spoken transcripts). Observing the grammatical correlates, Biber et al. (2004) classified these patterns into three major categories: NP/PP-based bundles, dependent clause bundles, and VP-based

bundles. Based on an analysis of the frequency of occurrence, they developed a structural framework for lexical bundles (see Table 1). As can be seen from the table, noun phrase (NP) and prepositional phrase (PP)-based bundles are composed of phrasal components. In many cases, these bundles are made up of noun phrase components, typically ending with a post-modifier (e.g., *the end of the*), or they consist of prepositional phrase components with embedded modifiers (e.g., *at the end of*) (Biber, 2006). Preposition phrase-based bundles begin with a preposition such as *after, as, at, by, in, for, of, on, to, over,* and *with*. Dependent clause bundles and VP-based bundles are similar in structure. Dependent clause bundles incorporate verb phrase elements just like VP-based bundles, but they also incorporate dependent clause fragments (e.g., *if we look at*) (Biber, 2006).

Table 1

Structural Taxonomy by Biber et al. (1999)

Structure	Examples
NP + *of*-phrase fragment	*one of the most, the context of the*
NP + other post-modifier fragment	*the fact that the, an increase in the*
PP + embedded *of*-phrase fragment	*in the case of, as a result of*
other PP fragment	*on the other hand, similar to that of*
anticipatory *it* + VP / adjective phrase	*it is necessary to, it is possible that*
passive verb + PP fragment	*is based on the, can be seen as*
copula *be* + noun phrase / adjective phrase	*is one of the, is similar to that*
(verb phrase +) *that*-clause fragment	*has been shown that, that there is a*
(verb / adjective +) *to*-clause fragment	*is like to be, has been shown to*
adverbial clause fragment	*as we have seen, if there is a*
pronoun / noun phrase + *be* (+…)	*there are a number, this is not the*
other expressions	*the presence or absence, as well as the*

Apart from the studies on major structural categories, considerable research has been carried out to identify and label the discourse functions of LBs. The culmination of

research has led to the development of functional taxonomies (Biber et al., 2004; Cortes, 2004; Hyland, 2008), which were generally influenced by research on systemic functional linguistics. In an effort to develop a taxonomy that can be applied to different registers, Biber et al. (2004) compared the discourse functions and frequency of LBs in classroom teaching and textbooks to conversation and academic prose. They described the discourse functions of LBs in three categories: stance expression, discourse organizers, and referential expressions. Table 2 illustrates that primary functions are sub-categorized according to their specific meanings and functions.

Table 2

Functional Taxonomy by Biber et al. (2004)

Category and Subcategories	Explanation and Examples
Stance expression	Convey attitudes/assessments of certainty of a given proposition (Biber, 2006)
A. Epistemic Stance	*It can be argued, the fact that the, are more likely to*
B. Attitudinal/Modality Stance	*it is possible to, it is necessary to, can be used to*
Discourse organizers	Reflect relationships between prior and coming discourse (Biber & Barbieri, 2007, p. 270)
A. Topic Introduction/Focus	*In this essay I, last but not least, in this section we*
B. Topic Elaboration/Clarification	*on the other hand, as well as the, has to do with the*
Referential expressions	Identify or frame particular abstract or physical attribute of an entity.
A. Identification/Focus	*those of you who, that's one of the*
B. Imprecision	*and things like that, or something like that*
C. Specification of Attributes	*in the case of, in the form of, as a result of*
D. Time/Place/Text Reference	*at the center of, the end of the*

The first category, stance bundles, has two meanings: epistemic and attitudinal/modal stance bundles. Epistemic stance is concerned with the knowledge status of the information. It expresses the degrees of certainty (e.g., *I don't know if*) or probability (e.g., *are more likely to*) of the following proposition. Attitudinal/Modality stance bundles express the speaker/writer's attitudes towards the actions or events described in the following proposition (Biber et al., 2004). These bundles reflect desire, obligation/directives, intention/prediction, and ability. For example, in the sentence "*Map analysis can be used to aid learning.*", "*can be used to*" is an attitudinal/modality bundle that expresses ability. It is important to note that stance bundles can be both personal and impersonal. Personal stance bundles are attributed to the speakers and are only used to express uncertainty while impersonal ones are used to express the degrees of certainty. The line of research by Biber and colleagues (Biber, 2006; Biber et al., 2004) suggested that most stance bundles are verb-based bundles.

Discourse organizers also have two categories. Topic introduction/focus signals the introduction and the focus of the topic (e.g., *if you look at, want to talk about, what do you think*). Topic elaboration/clarification bundles provide clarifications (e.g., *What I mean; has to do with*) and are used for comparisons and contrasts (e.g., *in contrast to the, on the other hand*). In Hyland's (2008) taxonomy, this group of functions correlates to text-oriented bundles. To develop a taxonomy, Hyland investigated the use of LB across four different fields, including both hard sciences and social sciences. According to Hyland, text-oriented bundles were more prevalent in the published articles in social sciences, pointing to the disciplinary differences between the two disciplinary fields.

Last but not least, reference expressions include four meanings. As Biber et al. (2004) describe, the identification/focus bundles are often used to help summarize or emphasize the main point after a lengthy explanation. Identifying/focusing bundles can also be used to introduce a discussion by stating the main points first and then giving details. In most cases, it has a discourse-organizing function, but it is classified under the referential expression category because the noun phrase following the identifying/focusing bundle indicates the groups of focus (Biber, 2006). As the name suggests, imprecision bundles indicate a reference that is ambiguous, unnecessary, or

additional references of the same type (Biber, 2006). Bundles specifying attributes help establishing logical relationships in a text and helps improve coherence. These bundles are used to specify an abstract or physical entity in terms of size, amount, or form. In addition, it can also identify abstract characteristics (e.g., *the source of the, the nature of the*). Lastly, Time/place/text-deixis bundles refers to specified time, place, or locations present in the text. Text deixis bundles make references to figures or graphs. In relation to structural categories, Biber et al. (2004) and Hyland (2008) have discovered that referential expressions were predominately noun- or preposition-based.

The creation of these taxonomies has brought a degree of standardization to the functional analysis of LBs and has resulted in a significant body of research focused on their usage in various genres and disciplines. However, it is also important to note that while some studies preferred Hyland (2008)'s taxonomy (e.g., Cooper, 2013), others adopted the taxonomy by Biber et al. (2004) (e.g., Chen & Baker, 2010). As argued by Ädel and Erman (2012), it is difficult to define functional categories of LBs objectively, and inconsistencies of categorical labeling across different studies reduce the generalizability of the findings. To illustrate, Chen and Baker categorize Focusing as Discourse organizing, whereas Biber et al. (2004) classify it as Referential (Ädel & Erman, 2012). Moreover, a category of LBs can exhibit multiple functions. As Chen (2009) exemplified, the lexical bundle *may be due to* is made up of two functional parts: "*may be*" provides information about the level of epistemic certainty whereas "*due to*" conveys an inferential assumption. According to Biber et al. (2004), a rule of thumb is to label these occurrences according to the most common uses found in concordance lines and corresponding contexts. Taking these issues into consideration, the next section discusses the literature on the differences in the use of LBs between non-native speakers and native speakers in academic writing.

2.4.2. Lexical bundles in academic prose

In the past few years, many lexical bundle studies have compared NS and NNS' production of lexical bundles in terms of the frequency of occurrence (tokens) and varied types of lexical bundles (i.e., varieties in grammatical structure and functional aspects). Although the focus has been on Chinese L1 users of L2 English (e.g.., Chen & Baker, 2010; Lu & Deng, 2019; Pan et al., 2016), other L1 learner groups have also been

investigated, such as Malay (e.g., Kashiha & Chan, 2015), Korean (e.g., Shin, 2018, 2019; Yoon & Choi, 2015), Japanese (e.g., Allen, 2010), Swedish (e.g., Ädel & Erman, 2012), Persian (e.g., Esfandiari & Barbary, 2017), and Turkish (e.g., Güngör & Uysal, 2016; Karabacak & Qin, 2013; Uçar, 2017). Regardless of the L1 background and academic register differences, most of the studies listed here used structural taxonomies proposed by Biber et al. (1999, 2004) and functional taxonomies by Hyland (2008) and Biber et al. (2004).

In the English as a second language (ESL) context, Chen and Baker (2010) compared 4-word lexical bundles in NNS and NS undergraduate student essays with that of expert native writers (i.e., published academic texts on hard science) focusing on structural and functional aspects. Regarding the structural categories, novice NNS and NS writers primarily relied on VP-based bundles (e.g., *as well as the*), which are typical of conversation register (see Biber, 1999). Expert native writers, on the other hand, were more likely to use NP-based and PP-based bundles. Compared to the other two groups, NNS primarily used VP-based discourse organizers, resulting in stylistically more wordy writings. Particularly, Chinese students relied on passive constructs with prepositions (e.g., *can be regarded as*). It is also worthy of note that NNS employed colloquial place/time deixis such as *all over the world* or other idiomatic constructs (e.g., in the long run) more than NS and expert NS academics, deviating from academic writing norms. Using a chi-squared test, Chen and Baker found that expert native writers used lexical bundles significantly more frequently and with greater variety than novice NS and NNS writers. Besides, the findings also showed that Chinese L2 learners tend to use bundles that are idiomatic and are used as connectors. According to Chen and Baker (2010), however, the results might be influenced by corpus size, as larger corpora tend to exhibit fewer bundles. Moreover, the genre differences across the corpora compared might also have influenced the findings.

Ädel and Erman (2012) extended the work of Chen and Baker (2010) by focusing only on 4-word lexical bundles in a single academic discipline to make the analysis more manageable. Adopting a comparable methodology, they examined the texts of native and Swedish NNS writers in the English as a foreign language (EFL) context. The findings of Ädel and Erman were consistent with Chen and Baker, supporting the idea that NNS

produce LBs less frequently and varied than their native counterparts. Moreover, similar to Chen and Baker (2010), lexical bundles in both NNS and NS groups contained high proportions of discourse organizers. Both studies showed that discourse organizers are common features of argumentative writing. Another striking finding was that the bundles with the initial preposition "*in*" differed across the two groups. Specifically, NS students mostly used in-bundles with abstract nouns (e.g., *in a variety of*), whereas in-bundles produced by NNS have concrete nouns (e.g., *essay, table, figure*) as complements. This might explain the overuse of discourse organizers in NNS writings. In addition, Ädel and Erman suggest that non-native users make use of evaluative/attitudinal bundles on a preposition (e.g., *it is easy to, it is difficult to*), which corresponds to stance bundles, to a greater extent than NS writers. Evaluative and attitudinal bundles are usually considered less inappropriate for academic writing as they signal subjective opinions. All in all, NNS learners tend to use clausal bundles (e.g., VP-based bundles) more frequently than NS. This tendency is also reflected in more recent research (e.g., Kashiha & Chan, 2015; Shin, 2019; Yoon & Choi, 2015). However, Shin (2018) asserts that the tendency of learners to use might be related to the nature of argumentative essays. Therefore, the high frequency of VP-based bundles cannot be solely attributed to the nature of NNS writing. Shin (2018) compared the LB use in Korean argumentative essays (approximately 1.6 million words) to a native speaker corpus (about 900.000 words) and found both corpora to include a high proportion of clausal-stance bundles.

The use of lexical bundles in doctorate dissertations and articles written by NNS and NS academic professionals have also received attention in the literature. Interestingly, several studies have found that non-native academic professionals employ a greater variety and frequency of lexical bundles in their writings than L1 writers (e.g., Güngör & Uysal, 2016; Lu & Deng, 2019; Pan et al., 2016; Uçar, 2017). However, the high frequency of lexical bundles might indicate overuse and a limited repertoire of LBs (see Granger, 1998). The bundles used by NNS academic professionals show similarities to NNS novice writers in terms of structure as both groups used more VP-based clausal bundles rather than PP-based or NP-based phrasal bundles. For example, a pioneering study published by Pan et al. (2016) examined 4-word lexical bundles produced by NNS and NS professional writers within a single discipline (i.e., telecommunications). They

found that NNS are more likely to use passive bundles (i.e., passive verb + prepositional phrase fragment), which are clausal in nature and VP-based. NS writers used the structure "*in the + Noun + of"*, which is a PP-based LB, with more than two times as many tokens as their non-native counterparts. The studies on an L1 background other than Chinese also reported the underuse of this phrase fragment (e.g., Chen & Baker, 2010; Esfandiari & Barbary, 2017; Lu & Deng, 2019; Shin, 2019; Yoon & Choi, 2015). The findings point to the heavy reliance on a limited number of framing bundles. Framing bundles, which are the subcategory of referential expressions, are commonly used in academic writing to present an argument in a cohesive way (Biber et al. 2004). Pan et al. (2016) suggested that these types of bundles can help readers focus on a given case (*in the case of*), highlighting aspects of an argument (*in terms of the*), or specifying the conditions (*in the context of*).

The line of research examining the NNS and NS use of LBs in academic prose provides insights into the general tendencies of L2 English writers from a variety of L1 backgrounds and offers evidence of NS norms in LB use. Overall, non-native writers were found to use fewer types and tokens of lexical bundles compared to NS writers. Also, NP-based and PP-based bundles are found to be underused while VP-based bundles are overused or used differently by NNS, especially novice writers. The literature comparing NNS novice and expert writers suggested that professional academic writers use lexical bundles more frequently in their writings. In other words, non-native speakers' use of lexical bundles becomes more frequent as they progress in their academic careers. However, the association between the overall frequency of LBs and language proficiency remains inconclusive. LBs across different proficiency levels require an examination to determine if there was a consistent pattern across all proficiency levels of non-native speakers. Therefore, the following section examines studies that demonstrate how lexical bundle usage varies across proficiency levels.

2.4.3. Writing performance and lexical bundles

An in-depth look at what needs to be learned is crucial to understanding the nature of Second Language Acquisition (Gass et al., 2014). As described by Gass et al., SLA "is the study of why most second language learners do not achieve the same degree of proficiency in a second language as they do in their native language" (p. 1). In this respect,

a great deal of scholarship has been dedicated to exploring the effects of linguistic features in L2 written responses on writing quality (see Crossley, 2020 for an overview). The literature has shown that the quality of writing performance correlates with various linguistic characteristics, such as cohesion (e.g., Crossley et al., 2016), syntactic complexity (e.g., Biber et al., 2011; Kyle, 2016), and lexical diversity (e.g., McNamara et al., 2010). In a longitudinal study, Crossley et al. (2016) explored the relationship between the production of cohesive and human holistic scores in descriptive essays of university students. Utilizing various cohesion measures (e.g., function word type-token ratio), Crossley et al. reported that four indices of cohesion explained 42% of the variance in overall essay quality scores. By leveraging the potential of automatic processing tools, it is possible to compute several indices related to syntactic complexity. Kyle and Crossley (2018), for example, compared three types of syntactic complexity indices related to clausal and phrasal complexity to assess their predictive validity on human ratings. The results indicated that fine-grained indices of phrasal complexity were the strongest predictor of writing performance as they accounted for approximately 19% of the variance in human ratings. They found higher-rated TOEFL independent essays to include more preposition objects, as well as adjectives and prepositional phrases that modify the preposition objects.

Previous literature has shown that there is a well-established link between writing performance and language features such as lexical complexity or syntactic complexity. However, phrasal complexity is a burgeoning, yet understudied area of research (see Bestgen & Granger, 2014). As Bestgen and Granger point out phraseological competence plays an important role in writing quality, and further work is needed in this domain of research. According to Flowerdew (2019), a high level of competency in lexical bundle use is the domain of only proficient English users. Lending support to this claim, Paquot (2018) compared the development of phraseological complexity in learner texts at different Common European Framework of Reference (CEFR) levels (B2, C1, and C2) with that of syntactic and lexical complexity measures. In a mixed effect model, measures of phrasal complexity (e.g., MI scores for adjective + noun collocations) explained the one-fourth of the variance in human judgements of writing quality for French EFL learners. Paquot demonstrated that phraseological competence measures can distinguish

between various CEFR levels effectively, particularly from B2 to C2. In another study, Garner et al. (2019) analyzed how n-gram indices relate to holistic human in L1 Korean learners of English argumentative essays. Specifically focusing on trigram and bi-gram measures, they found phraseological complexity to account for one-fifth of the variance in essay scores. The frequency of academic bi-gram use and essay scores positively correlated. It was also shown that learners at high proficiency levels use more non-finite clauses, more adverbials, and more adverbial prepositions than low-proficient learners (Kyle, 2016). Using TAASSC (Kyle, 2016), Kyle found that indices of phrasal complexity explained 20% of the variance in essay scores. Besides token frequency, more proficient learners used more complex and diverse phrases, particularly prepositional phrases and adjectival modifiers. Evidence from a growing body of research showed that phrasal complexity is an effective predictor of L2 writing performance similar to syntactic complexity and lexical complexity. However, some studies found no significant relationship between writing quality and the use of formulaic language (e.g., Kılıç, 2015; Torlak, 2020). Torlak, for example, examined the relationship between essays (opinion and problem-solution essays) written by Turkish EFL students in an exam condition and overall writing quality. Interestingly, the study did not reveal any significant results indicating a correlation between human holistic scores and formulaic language use.

Phraseological complexity can predict writing scores, but confirming the link between lexical bundle use and writing quality presents a challenge. Proficiency is measured or defined in various ways: by years of study or educational level (e.g., Ruan, 2017; Qin, 2014), by standardized tests such as IELTS (Appel & Wood, 2016; Cooper, 2013; Staples et al., 2013), by assigned levels by institutions (Hazeyaaji, 2022), or by other human holistic scores (either with or without the assistance of automated scoring) according to the proficiency levels described in the CEFR (Appel, 2022; Chen & Baker, 2016; Kim & Kessler, 2022; Shin, 2018; Vo, 2019). Many researchers have investigated the relationship between the use of lexical bundles and writing quality/performance, complementing the broader studies of phraseological complexity. Staples et al. (2013), for example, examined the use of LBs in students' responses to the TOEFL iBT test. The total corpus size was 249.417 words from 480 participants from different nationalities, and they subdivided the corpus into three proficiency levels (low, medium, and high

proficiency) based on ETS scores given to two written tasks (integrated and independent writing tasks). The findings indicated that there was a decrease in the frequency of bundles used as the proficiency level increased, with high-scoring essays having the least amount of bundle tokens. Low-scoring students used prompt-related bundles more than high-scoring learners. Staples et al. suggested that learners at lower levels largely rely on formulaic devices, and high frequency of LBs is a feature of low-proficiency learners. Moreover, Staples et al. reported that the number of stance bundles followed a similar distribution across all proficiency levels. The overreliance on stance expressions might be related to the nature of argumentative or opinion essays. Independent writing tasks require the learner's opinion on a particular topic, and this might lead to the production of personal (e.g., *in my opinion, I agree with the*) and impersonal stance expressions (e.g., *it is important to*). This might also explain why VP-based bundles are prevalent in NNS argumentative writing because such LBs as "*in my opinion*" are clausal in nature (i.e., PP as adverbial + the main clause) and are made up of commonly found words.

The corpora Staples et al. (2013) used were comparable in terms of size; however, as argued by Chen and Baker (2016), they failed to control the variables of L1 background and task type. Chen and Baker extended Staples et al. by examining the rated (B1, B2, and C1 in CEFR) expository and argumentative essays of L1 Chinese learners of English by controlling task and L1-related confounding variables. The results showed C1 level learners used more PP-based bundles and fewer VP-based bundles than the other two groups. In more detail, all the PP-based bundles in B1 and half of the bundles in B2 were other PP fragments, which mostly contained adverbial phrases without -of fragments. Regarding the NP-based bundles, B2 groups were found to use the highest proportion of NP-based lexical bundles among other groups, but their use of NP-based bundles was not typical of academic writing (e.g., *more and more people*). Moreover, learners at higher proficiency levels used fewer conversational or speech-like bundles. The decreasing use of conversational or colloquial bundles at higher proficiency levels indicates that they become more aware of the difference between formal and informal at higher levels.

In a testing context, Kim and Kessler (2022) analyzed the use of 3-, 4-, and 5-word LBs in argumentative essays written by Chinese L2 learners at different proficiency levels. The essays were written in a timed manner without consulting additional reference

materials. The resulting corpus contained approximately 18.000 words with an average essay length of around 150 words. Overall, the 3-word bundles were the most common and low-scoring groups employed these bundles more frequently than high-scoring ones, but their use of bundles was rather colloquial and inappropriate. However, while low-scoring essays contained more instances of 3-word bundles, high-scoring essays exhibited more types of these bundles. For 4-word PP-based lexical bundles, students frequently used the expressions, *with the development of, at the same time* and *on the other hand*. The high frequency of transitive constructs might be related to the nature of timed essays (see Paquot, 2010). In addition, the high-scoring learners in Kim and Kessler (2022) frequently used prompt-based bundles, making their writings more dependent on the content of the prompt.

Another recent study by Appel (2022) examined lexical bundles in differing lengths (3 to 7 words LBs) in argumentative writing untimed assignments of L1 Japanese English learners using two sub-corpora (high and low scoring learners), which contained essays scored by human raters. In Appel, the student wrote the essays in untimed task conditions, and they had access to reference tools for checking their accuracy of use. The results of Appel suggested that the tokens and types of lexical bundles increased with proficiency levels, particularly three-word lexical bundles. Therefore, Appel's findings supported the previous research (e.g., Huang, 2015; Kim & Kessler, 2022; Vo, 2019) that found a positive correlation between lexical bundle frequency and proficiency level. Also supporting previous studies (e.g., Staples et al., 2013; Vo, 2019), low-scoring essays consisted of more prompt-influenced bundles than high-scoring ones.

From a different perspective of proficiency, Ruan (2017) investigated 4-word lexical bundles in Chinese learners of English at different years of study or educational levels (first-year to fourth-year dissertations). Ruan divided the corpus (approximately 1.2 million words) into four sub-corpora: Y1 First Essays, Y1 Final Coursework, Y2 Final Coursework, and Y4 Final Year Dissertations. Interestingly, the findings suggested that as learners progressed to higher educational levels, they used a greater variety of lexical bundles but with decreasing frequency. Y1 First Essays had a much higher average token frequency of 199.3 compared to Y4 Final Year Dissertations, which had an average token frequency of 49.2. The varied use of bundle types supports Vo (2019), Appel (2022), and

Kim and Kessler (2022). However, the findings on the frequency of the tokens might be affected by the effect of textual composition (e.g., essay length, the number of essays) in the corpora compared. While Y1 First Essays corpora consisted of 90.729 words and had an average text length of 396 Y4 Final Year Dissertations had about 475.000 words and an average text length exceeding 8000. The comparisons across corpora with different corpus sizes might be unreliable and impact the representativeness of LBs found even with the normalization procedures (Hyland, 2012). As Hyland puts it, smaller corpora produce more lexical bundles than larger ones as smaller corpora utilize much lower frequency cut-off points. In terms of structural differences, as learners progressed through higher levels of education, the use of PP-based LBs increased. Therefore, students acquire VP-based and NP-based bundles before PP-based bundles (Ruan, 2017). This view is also supported by Vo (2019), who found a similar pattern of structural distribution of LBs.

Overall, the research concerning the link between writing quality and the use of lexical bundles produced mixed and uncertain results. While some studies suggested that learners at higher proficiency levels use lexical bundles with greater variety and frequency, others reported contradictory findings. As noted earlier, only a few previous studies have identified LBs in L2 argumentative essays, often without comparable L1 data (e.g., Staples et al., 2013). Additionally, there is inconsistency across studies in how proficiency is defined. As can be seen from the literature, learner proficiency was determined in various ways such as by standardized tests (e.g., Appel, 2022) and educational level (e.g., Ruan, 2017). While the diversity of definitions to describe proficiency draws a more comprehensive picture of the relationship between proficiency and LB use, it limits the comparability of findings across different studies. Nevertheless, the body of research investigating the relationship between proficiency and LB use had valuable insights into learners' use of LBs. For example, the majority of the studies have shown that learners at lower proficiency levels use more conversational lexical bundles than learners at higher proficiency levels (e.g., Chen & Baker, 2016; Kim & Kessler, 2022). In addition, the studies investigated suggested that as learners advance in proficiency, their essays contain more types and tokens of PP-based bundles. For example, learners at higher proficiency levels can use English prepositions beyond their adverbial meanings (e.g., *in my opinion I*, *in my opinion the*), making their writings more

complex in terms of structure. However, the link between writing performance and lexical bundle frequency should be approached with caution as there is a limited number of research conducted using statistical analyzes such as linear regression (e.g., Candarlı, 2020).

2.4.4. Accuracy studies on lexical bundles

Contrastive Interlanguage Analysis (CIA) (Granger, 1996, 1998, 2015a) has been widely practiced as an effective model for corpus research on learner language to identify learner deviation (i.e., overuse and underuse) from target norms. Previous corpus-based studies that compared NS and NNS lexical bundle use have shown that non-natives use fewer and/or less varied bundles. While some studies suggested that the underuse of LBs might be due to a lack of knowledge or instruction (e.g., Chen & Baker, 2010), others claimed that it is related to insufficient exposure to authentic language use (e.g., Staples et al., 2013). Additionally, Schmitt and Carter (2004, p. 13) cited several studies that suggest linguistic sequence overuse could be related to "the tendency to stick with familiar and safe sequences". Certain native-speaker-like phrases such as *on the other hand* are found to be frequently used in NNS corpora (e.g., Chen & Baker, 2010; Shin, 2019), possibly due to the learners' familiarity with these structures (see Hasselgård, 2019).

There is no doubt that L2 English learners have difficulty using phraseological patterns in their writing (see Flowerdew, 2019; Paquot & Granger, 2012; Schmitt & Carter, 2004), but very few studies have examined LB accuracy. To the authors' knowledge, only a handful of studies investigated the errors in LBs (e.g., Bychkovska & Lee, 2017; Geluso, 2022; Huang, 2015; Lee et al., 2020; Shin et al., 2018; Uzun, 2018). Huang (2015), for example, shed light on the most common errors in 3-4-5-word LB use and the distribution of errors identified from argumentative essays of junior and senior Chinese EFL majors. Huang used a Mann-Whitney U Test ($p < .05$) to compare mean accuracy scores to find a statistical difference between the two groups. Surprisingly, the study reported that senior Chinese EFL learners made as many errors as their junior counterparts despite being more experienced (with over 90% accuracy rate). A similar study was conducted by Bychkovska and Lee (2017) on the accuracy of lexical bundles in essays produced by Chinese ESL learners. The findings of Bychkovska and Lee (2017)

contradicted Huang (2015) as they found that 77% of the total bundles were accurate, and approximately half of the inaccurate uses were related to misuse or omission of articles and prepositions. Notably, the PP-based bundles *on the other hand* and *at the same time* were among the most misused bundles.

Shin et al. (2018) analyzed the uses of definite articles embedded in four-word lexical bundles in the argumentative essays produced by L2 Korean learners. As opposed to the widely held belief that NNSs use LBs less frequently and with less variation, Shin et al. suggested that misuse of lexical bundles may contribute to their underuse since frequency-based corpus analyzes do not recognize incorrectly used LBs. For example, *in the other hand* is not recognized as a lexical bundle in n-gram searches. Consequently, such misformed uses are not added to the overall frequency counts of the LBs. To address this issue, they analyzed the lexical bundles using an innovative approach called core expression analysis. A core expression is a phrase that forms around a core word, which carries the main meaning of the lexical bundle; for example, for the bundle *in such a way*, the core expression is *such a way*. To extract core expressions from lexical bundles containing definite articles, the authors removed adjoining articles that surround the core word. To give an example, for the bundle *on the other hand*, the core expression is *other hand*. Shin et al. manually checked the words to the left and right of each use of a core expression to explore the uses of core expressions in context. Specifically, the study analyzed the concordance lines including at least two to three sentences on either side of the core expressions to investigate the definite article use. The study then generated a list of core expressions and added those core expressions to overall LB counts. The findings revealed that there were significantly more core expressions (77 types and 2800 tokens) than four-word lexical bundles (41 types and 428 tokens) in the learner corpus. Consequently, the analysis of core expressions revealed that L2 learners used more LBs than those reported in traditional frequency-based analyzes of articles. Replicating the study of Shin et al. (2018), Uzun (2018) analyzed the erroneous lexical bundles in the Turkish context to test the functionality of the core expression approach. Uzun reported similar findings, as all error types were related to the omission of articles. The Turkish language share similarities to Korean in that it does not have an article system, so the omission of articles was not unexpected.

As an extension to Shin et al. (2018), Lee et al. (2020) investigated the use of PP-based bundles in terms of overuse, underuse, and misuse in a corpus compiled of Korean students' argumentative essays. Lee et al. adopted a modified version of core expression analysis as many of the core expressions in PP-based bundles include a content word that is either preceded or followed by a preposition. To address this problem, they focused only on the accuracy of initial prepositions in LBs. For example, they searched for the core expression *front of* from the bundle *in front of my* to analyze the use of the preceding preposition "*in*". Despite modifications to the core expression approach, the analysis ignored prepositions found in the middle and at the end of lexical bundles. The accuracy analysis showed that the total error rate was quite low, at 7.13%, partly due to the underuse and the lack of variety of PP-based LBs (Lee et al., 2020). Most errors in Lee et al. were related to the misuse of prepositions (70% of the total preposition errors), PP-based LBs with *in*, *for*, *on*, and *at*. Thus, most of the core expressions of LBs were correctly used, but they were often combined with incorrect prepositions.

In conclusion, Contrastive Interlanguage Analysis (CIA) has proven to be an effective model for corpus research on learner language to identify learner deviation from target norms. Studies that have investigated LB accuracy indicate that L2 English learners have difficulty using phraseological patterns in their writing, but it is evident that very few studies have examined LB accuracy. Moreover, prepositions are important constructs for academic writing, and a significant body of research has shown that learners commonly struggle with them. Considering the fact that Turkish is a postpositional language (Underhill, 1976), and the use of prepositions poses a challenge for Turkish students (see Demirel, 2017), this study aims to analyze the errors in four-word PP-based bundles. Adopting a methodology similar to core expression analysis, the present study aims to test the function of regular expression codes for analyzing the erroneous use of lexical bundles.

CHAPTER 3: CORPORA AND METHODOLOGY

This exploratory study investigated the use of preposition phrase-based lexical bundles. To increase the generalizability of formulaic language studies, Paquot and Granger (2012) state that there is a need for explicitly documented data, which gives comprehensive details about learner characteristics, setting (e.g., frequency and dispersion thresholds), and classification criteria applied. Therefore, this chapter gives information about the corpora used, lexical bundle extraction methods, statistical tests used, and data analysis procedure.

3.1. Data

3.1.1. Corpora

Data of this study consists of two corpora: the Louvain Corpus of Native English Essays (LOCNESS), and Turkish English Learner Corpus (TELC) students. Each corpus contains writing samples of essays collected employing a cross-sectional design in corpus compilation. Preparing corpora for data analysis involved a three-dimensional procedure: (1) gathering and selecting the texts for the corpora, (2) cleaning unnecessary information, and (3) categorizing the data according to exclusion criteria.

3.1.1.1. Non-native corpus

[1]TELC is a publicly available cross-sectional learner corpus compiled of written essays by Turkish EFL students from four different universities in Türkiye. It has been collaboratively created by a research team working on a project (Project No: 220K289) supported by TÜBİTAK (Research Council of Turkey) 3501 Grant. The first step of building the corpus involved adding the participants to a Google Classroom and asking them to fill in the informed consent forms. After that, students (n = 262) received instruction to write an opinion essay in response to a given writing prompt under time constraints (1 hour) in a classroom environment. The students were allowed to use

[1] TELC is publicly accessible at Telcorpus.com

reference tools such as online resources and dictionaries, but the range of reference tools available were limited in order to maintain authenticity of the data. The writing prompt consisted of four topics (banning cars, hybrid education model, empathy, and the internet), and students were free to respond to either one or multiple texts by giving their opinion. Also, the essays were written in a manner that did not require learners to synthesize information from reading or listening sources. Therefore, all the essays were independent essays. Below are the given writing prompts:

1. Cars should be banned from city centres to reduce traffic problems in big cities. Do you agree or disagree? To what extent do you agree? Explain your reasons using examples.
2. Universities should adopt a hybrid education model instead of online education. Do you agree or disagree? To what extent do you agree? Explain your reasons with detailed examples.
3. Empathy is considered to be one of the essential personal/social skills in the 21st century. Do you agree or disagree/To what extent do you agree? Explain your reasons using examples.
4. The Internet has caused people to be isolated from their real lives. Do you agree or disagree? To what extent do you agree? Explain your reasons using examples.

Students submitted a total number of 691 essays. The collected essays were checked against plagiarism and converted to UTF-8 encoding because it is compatible with most corpus tools. The next step was to manually remove unnecessary parts of the text (i.e., title, footers, reference lists) and duplicate texts. According to Pan et al. (2020), LB extraction is affected by essay length, so short essays (fewer than 50 words) were excluded. The final version of TELC consisted of 192.003 words and contained 687 essays with an average length of 276 words, the shortest essay having a word count of 92 words, and the longest one having 685 words.

To give additional information on the profile of participants, The age of the students ranged between 18 to 47, with an average age of 20. Most of the participants were female (f = 420, % = 63), and approximately one-third of the participants (f = 420, % = 63) were male. It is assumed that most participants have had at least ten years of

instruction in English in a foreign language context since students take English as a compulsory subject nationwide from primary school until the end of secondary school. Also, enrollment in university programs requires students to take the national-level university examination and get a passing score. Administrated by the Measuring, Selection and Placement Center (ÖSYM), the national language examination is part of a standardized university entrance exam which includes 80 multiple-choice questions in English focusing on translation ability, reading comprehension, vocabulary, and grammar. Though this high-stakes language test in Türkiye lacks the balance of all four language skills (speaking, listening, writing, reading), it proves that the participants in this study are proficient enough to enrol in an English-related university program. The students in TELC have taken English preparatory classes or have majored in an English-related field (e.g., English literature, English language teaching) at four universities in Türkiye: İstanbul University, Trakya University, Ankara University, and TED University. Overall, the students are assumed to be at B2 level as they received 600 hours of B2-level instruction covering all four skills of English, as well as grammar use.

3.1.1.2. Native corpus

[2]LOCNESS is compiled of monolingual British and American students' written argumentative and literary essays, which includes 324.304 words in total and contain both literary and argumentative essays. All students in this corpus are fully English native speakers and they received at least "upper-second class" grade. Since, LOCNESS is a compilation of essays written by university students, it represents novice academic writing. The corpus is compiled of different types of essays: examination papers, timed essays, and untimed essays on topics such as transport, environment protection, education, and arts. Also, all essays are independent essays (i.e., without synthesizing information from reading or listening).

To make more direct comparisons with TELC, the present study only used argumentative essays in LOCNESS. Moreover, to ensure accurate quantitative analysis, unnecessary elements such as headings, brackets indicating quotations, and other

[2]The Center for English Corpus Linguistics (CECL) at Université catholique de Louvain provided access to the LOCNESS corpus, which was instrumental in conducting this research. LOCNESS is accessible at https://www.learnercorpusassociation.org/resources/tools/locness-corpus/

irrelevant content were removed from the corpus. The final version of LOCNESS contained 229.059 words and contained 322 essays with an average length of approximately 709 words; the shortest essay was 179 words long, whereas the longest essay was 2925 words long. The essay selected for this study are 114 British A-level student essays, 175 American university student essays, and 33 essays from British university students. Failing to control these variables can affect the results of the study as textual composition and genre differences act as confounding variables. The reason for selecting LOCNESS as a reference corpus was to control the important variables such as:

1. corpus size (small corpora)
2. the population of students (undergraduate and first-year graduate students)
3. genre (i.e., argumentative essays)
4. essay type (independent)

3.1.1.3. Normalization

As can be seen from Table 3, TELC contains more essays, and the essays are relatively shorter compared to LOCNESS. Textual composition (i.e., corpus size, average text length, and the number of essays) of the corpora compared can affect the results of LB token counts (Biber et al., 1998). For example, a corpus with lengthier essays might yield a larger number of 4-word LBs than a smaller corpus. The current study normalized the raw frequencies found in TELC and LOCNESS to ensure the accuracy of LB type/token counts and to guard against the potential inflation of the results.

Table 3

Textual Composition of TELC and LOCNESS

Corpus	Number of essays	Average text length	Corpus size
TELC	687	276	192.003
LOCNESS_argumentative	322	709	229.059

Note. LOCNESS_argumentative *is hereafter referred to as LOCNESS.*

Normalization is a technique frequently employed in corpus-based investigations to adjust raw frequency counts from texts of different lengths or corpora of different sizes

(Biber et al., 1998). By taking account of the total number of words, the raw frequency counts should be divided by the number of words in the text and multiplied on a chosen basis criterion. Frequency counts should also be normed to the typical text length in a corpus as the higher basis of norming results in inflated frequency counts and decreases the reliability of comparisons. For this study, the basis for normalization was chosen as 500 since essays in both corpora are about this length in word count. The normalization of raw frequencies for each essay in TELC was calculated with the following formula:

$$\text{Normalized Frequency} = \frac{\text{LexicalBundleTokens}}{\text{EssayWordcount}} \text{ x } 500$$

3.1.2. List of reference bundles

According to Cortes (2013), corpus-driven identification of LBs requires a corpus containing a minimum of one million words. It implies that the corpus used in this study is too small to extract lexical bundles in a corpus-driven way, so using an empirically derived list might be a good alternative. Therefore, the present study used Biber et al.'s (1999, 2004) list of four-word PP-based LBs as reference bundles. Biber et al. generated this reference list of LBs from *Longman Spoken and Written English* corpus (LSWE; Biber et al., 1999). *LSWE* is a native speaker corpus consisting of over 40 million words in registers, such as conversation, fiction, newspaper language, and academic writing. Lexical bundles in academic prose were studied using the academic prose sub-corpus of *LSWE* (roughly 5,3 million words), which comprises research articles and advanced academic books across many disciplines. Since argumentative essays are in the domain of academic writing, Biber et al.'s list can serve as a proxy for the typical use of lexical bundles in academic prose. In addition, a corpus-derived lexical bundle list can be affected by topic variation as the topics in LOCNESS and TELC are not strictly matched. Therefore, using a reference list of bundles can control the prevalence of context-dependent bundles and ensure the reliability of the analysis. For example, Vo (2019) conducted a corpus-driven analysis and found writing prompts to cause topic effect, which inflated the frequency counts of NP-based bundles in low-proficient learners.

The current study only investigated 4-word PP-based LBs found in Biber et al. (1999) as "reference bundles", most of which are phrasal in nature. Preposition phrase-

based bundles are groups of words that begin with a preposition and end with a noun, pronoun, or noun phrase and these word combinations consist of a preposition, its object, and any words that modify the object. To illustrate, Biber et al. (1999) classified those bundles into two categories: prepositional phrases or prepositional phrase fragments with an of following the noun, and other prepositional phrases or prepositional phrase fragments (see Table 4). According to Biber et al., when a preposition phrase is written alone, it is identified as a PP fragment**.**

Table 4

List of Reference PP-Based Bundles in Biber et al. (1999)

PP with of-phrase fragment	about the nature of, as a result of, as a function of, as part of the, as a consequence of, as a matter of , as a means of, as part of a, at the end of, at the time of, at the beginning of, at the level of, at the expense of, to the start of, by the end of, by the presence of, by the end of, for the development of, for the purpose of, for the purposes of, from the point of, in the case of, in the absence of, in the form of, in the presence of, in a number of, in terms of the, in the context of, in the course of, in the development of, in the number of, in the process of, in a variety of, in the area of, in the direction of, in the face of, in the formation of, in the pathogenesis of, in the study of, in the treatment of, in the use of, in view of the, of a number of, of some of the, of the effects of, of the nature of, of the use of, on the basis of, on the part of, on the surface of, over a period of, to the development of, to that of the, to the presence of, to the use of, with the exception of, of the nineteenth century
Other PP fragments	as in the case, at the same time, between the two groups, by the fact that, for the first time, from the fact that, in such a way, in the same way, in the present study, in a way that, in addition to the, in an attempt to, in contrast to the, in relation to the, in the early stages, in the first place, in the next chapter, in the next section, in the sense that, in this case the, of the fact that, of the most important, , on the other hand, on the one hand, on the grounds that, similar to that of, similar to those of, to the extent that, to the fact that, with respect to the, in the United States, in England and Wales, in the United Kingdom, in the nineteenth century

Prepositional phrase with embedded of-phrase fragment refers to phrases with "*of*" particle functioning as post-modifier of a noun. There are also prepositional phrases

without an embedded of-phrase that are classified as other prepositional phrases. As can be seen from the Table above, most reference bundles (35 bundles) have the initial preposition "*in*", which makes it the most common initial preposition in PP-based bundles. To mitigate the effect of contextual differences, context dependent bundles such as "*in the United Kingdom*", "*of the nineteenth century*", "*in the nineteenth century*", "*in England and Wales*", "*in the United States*" were manually excluded from the list of reference bundles. In the final list a total of 86 were considered as reference bundles.

3.1.3. Human holistic scores

In some cases, when L1 and L2 groups are compared, the identified production patterns may not be related to the L1 background but rather result from differences in proficiency levels (Appel & Lewis, 2020). Therefore, similarities or differences in the use of PP-based bundles in NS and NNS corpora can be related to the nature of non-native writing. A comparison between low-scoring and high-scoring essays can offer a more comprehensive understanding of the nature of learner language.

The rubric for grading writing quality is the joint work of the project team. A group of experts in the field developed a scoring rubric to standardize and increase the reliability of the essay scoring procedure by adopting the Delphi Method (Helmer & Dalkey, 1963). The resulting rubric included categories of assessment such as vocabulary, grammar, task completion, coherence and cohesion, and spelling and punctuation - categories that are commonly used in standardized tests such as IELTS and TOEFL (see Appendix 2). The two raters graded the essays and a third rater intervened in case of disagreement ($\kappa = 0.75$, indicating a strong agreement). The scores for each essay ranged on a scale from 1 to 10, with the score of 10 being the highest score.

3.2. Data extraction and exclusion

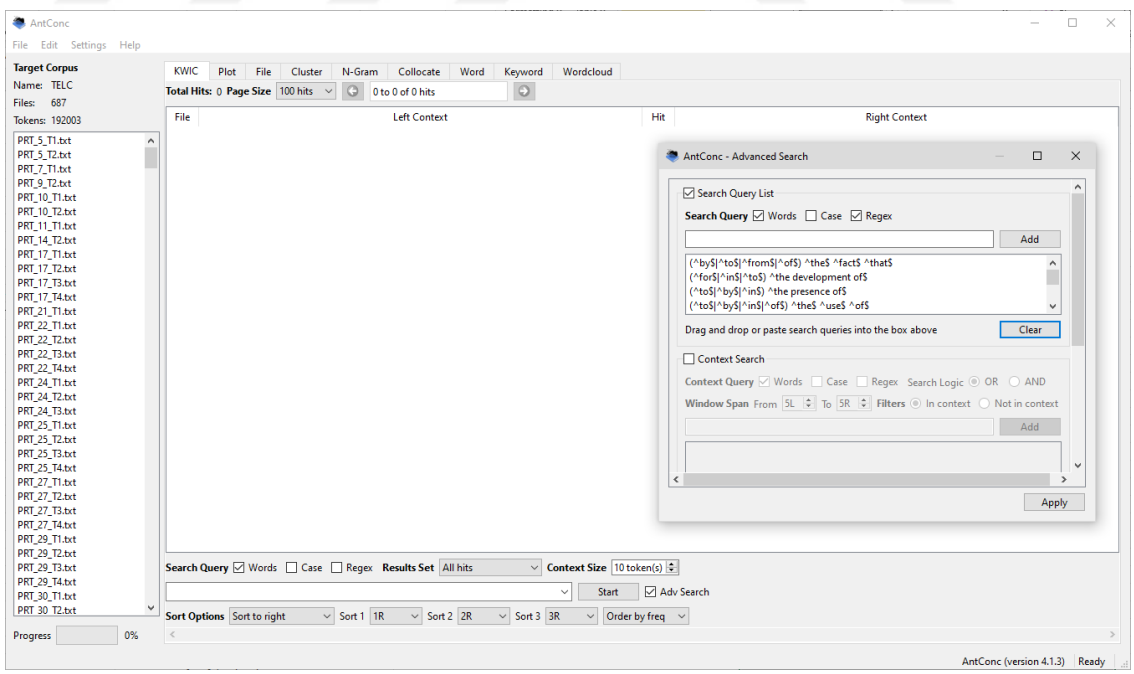
This section outlines the data analysis procedure, as well as the tools that are used to extract lexical bundles and regular expressions. Additionally, it provides an overview of the error classification taxonomy.

3.2.1. Instruments for extraction and analysis

Formulaic language has been extracted from learner corpora using many different techniques, all of which involve some level of automation (Paquot & Granger, 2012). AntConc 4.1.4 (Anthony, 2022) is one of the corpus tools that can automatically identify and extract lexical bundles. Before loading the corpora to AntConc, the texts were manually checked for excluding irrelevant content such as figures, titles, page numbers, participant ID numbers, and XML tags. As a next step, the texts were uploaded to Antconc and the "Search Query List" section was used to extract four-word reference bundles from NNS and NS corpora (see Figure 2), as it allows batch extraction of lexical bundles. The extracted bundles in each corpus were manually checked for cases misidentified by AntConc, for example, due to punctuation and contractions.

Figure 2

Screenshot of N-Gram Display in AntConc (Anthony, 2022)



3.2.2. Lexical bundle extraction

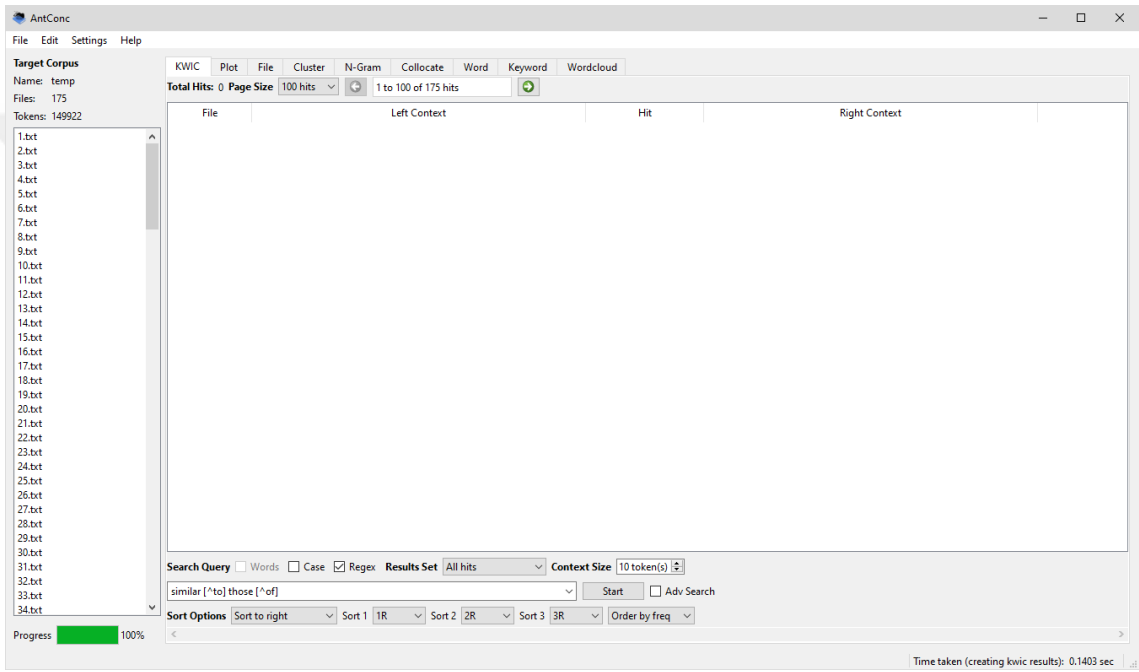
A vast majority of lexical bundles in academic prose are not complete structures but are parts of noun phrases and prepositional phrases, such *in a number of* and *as a result of*. The length of lexical bundles ranges between three words to seven words. To limit the scope of the current study only focused on 4-word lexical bundles because three-word lexical bundles are too prevalent and harder to interpret (Csomay, 2013). Since lexical bundles longer than four words are quite rare (Biber et al., 1999), they are less likely to be extracted. This is especially true for smaller corpora such as TELC and LOCNESS, which contains learner essays. Therefore, analyzing longer bundles might yield insignificant results. In addition, four-word sequences are found to be the most researched length for writing studies (e.g., Chen & Baker, 2010; Cortes, 2004; Lee et al., 2020; Shin et al., 2018; Uzun, 2018), probably because the number of 4-word bundles is often within a manageable size for manual categorization (e.g., error tagging) and concordance checks.

3.2.3. Extraction of attempted bundles

Since prepositions are common components of core expressions, the extraction of core expressions from PP-based lexical bundles requires a modified core expression analysis (Lee et al., 2020). This modified analysis involved conducting multiple searches to ensure accurate extraction of core expressions from prepositional phrase-based lexical bundles. As a result, the analysis of prepositions embedded in LBs using core expression analysis necessitates additional effort to extract every relevant bundle (Lee et al., 2020). To address this issue, similar to the list of core expressions given by Lee et al. (2020), a list of regular expressions (see Appendix 1) was used to extract the erroneous uses of lexical bundles for analysis. To extract the regular expression, the "Regex" function of Antconc was used (see Figure 3). Since the present study investigates PP-based lexical bundles, errors related to other structures (e.g., other function words, content words, etc.) in bundles are also considered.

Figure 3

Regex Function of Antconc (Anthony, 2022)



Regular expression codes used can account for the lexical bundles with double prepositions or other function words. To illustrate, the regular expression for "*at the beginning of*" in TELC, the following codes were used: "[^at] (\w|\W) ^beginning$ ^of$" and "(^at$) (\w|\W) ^beginning$ [^of]"; however, the extracted bundles differ from the targeted reference bundles. The query of regular expression for *at the beginning of* resulted in such bundles as: *at the beginning*, *by the beginning of*, and *from the beginning of*. The use of regular expressions can save time and manual effort as the set of regular expression codes can be copied/pasted to search queries for batch analysis of all regular expressions by using the "Adv Search" option in Antconc software. It can also reduce the inflation of retrieved results since the analysis excludes correct uses of lexical bundles. After retrieving the regular expressions in the learner corpus using AntConc

software, the regular expressions were manually checked for the search of attempted bundles. Overall, from 86 PP-based reference bundles, a list of 119 regular expression codes was created for the analysis.

3.2.4. Error classification taxonomy

Even though LB extraction is an automatic process, the extracted bundles are still raw materials that require manual interpretation and analysis of errors. One of the objectives of the current research is scrutinizing TELC for learner errors, as a thorough investigation of learner corpus can provide insights into the accuracy of use and intelligibility of the learner texts. As stated by Dulay et al. (1982), "Learners may omit necessary items or add unnecessary ones; they may misform items or misorder them" (p. 151). Based on Dulay et al., the present study classified errors into three categories: omission, addition, and misformation.

Omission errors are characterized by the absence of a grammatically required element in a specific position within a sentence, such as using *is based the* instead of *is based on the*. Addition errors involve unnecessary inclusion of an element within a sentence, for example using *in the terms of the* instead of *in terms of the*. Misinformation errors refer to the use of wrong forms within a sentence, such as using *with the same time* instead of *at the same time*. The errors were classified and coded by two raters using the error classification framework by Dulay et al. (1982). To address the potential for human error and personal bias in error analysis, this study conducted a Kappa Coefficient test to assess inter-rater reliability. The first rater was the author of this study, while the second rater was a native English speaker who volunteered to code the examples via a Google Forms Survey. Beforehand the rating process, the volunteer rater received instruction on the error categories given by Dulay et al. (1982) to ensure consistency in the coding process. In cases of disagreement, the two raters negotiated to reach a full agreement and asked for another native speaker's opinion to help resolve the issue. To evaluate the interrater reliability a Cohen's Kappa was conducted. Kappa (κ) value ranges from 0 to 1, and the closer it is to 1, the level of agreement between the raters increases. The resulting value was $\kappa = 765$, with a p-value of $< .001$, indicating a high level of agreement between the two raters.

3.3. Data analysis procedure and statistical tests

This section explains the procedure of data analysis in detail. To address the three research questions, quantitative corpus-based analyses were employed. The resulting data were statistically analyzed using both parametric and non-parametric statistical tests to compare means. To follow the established conventions, for all statistical tests, *p* value was set to $\alpha = 0.05$ confidence levels to reduce the chance of type I error (i.e., false positives). Table 5 provides a summary of the data analysis procedure.

Table 5

Summary of the Data Analysis Procedure

Research Questions	**Instruments**	**Data Analysis Procedure**
RQ1: Is there a difference between L1 Turkish L2 English students' use of 4-word PP-based bundles compared to those produced by native speakers of English?	TELC and LOCNESS Reference bundles list Regex codes for correct uses Antconc v4.2	Retrieve LB types and tokens in LOCNESS and TELC. Compare frequencies across two corpora using T-test. Identify the overuse and underuse of lexical bundles using Mann-Whitney U test.
RQ2: How accurately do L1 Turkish L2 English students use PP-based lexical bundles? How do they use the regular expressions of each lexical bundle?	TELC Regex codes for deviations Error taxonomy	Identify and classify the errors in regular expressions. Conduct a Kappa Coefficient test for interrater reliability. Add attempted bundles to overall lexical bundle counts.
RQ3a: Does 4-word PP-based bundle frequency predict the writing performance of Turkish students? RQ3b: Does the frequency of 4-word PP-based differ between high-scoring and low-scoring essays?	TELC AntConc v4.2	Retrieve LB types and tokens for high-scored and low-scored essays. Investigate the association between LB use and human holistic scores. Conduct a linear regression test and identify the predicting value of LB use.

3.3.1. Addressing research question 1

The first question aims to investigate the use of PP-based reference bundles in LOCNESS and TELC in terms of LB type and token distribution. A quantitative design was adopted to examine the extent of differences between English-speaking L1 Turkish and native English speakers in the use of 4-word PP-based bundles. A reference bundle list of preposition phrase bundles derived from Biber et al. (1999) was used to identify bundles. Based on this list, both corpora were searched for the 86 reference bundles using regular expression codes for correct uses.

The results of the normalized frequencies across the two corpora were evaluated using independent samples t-tests whenever the assumptions of normal distribution and equality of variances were met, and the non-parametric version of the t-test, the Mann-Whitney U test, was used if the assumption of normality was violated. The normality of the data was verified by kurtosis and skewness values. In the case of values above or below 1.50, excess z-scores above ±3 were excluded from the analysis to exclude the outliers. The data did not follow a normal distribution, so four samples were removed from the analysis to conduct an independent samples Student's t-test to compare two corpora. By analyzing the results of the statistical tests, it was determined whether the total bundle tokens in TELC occur excessively frequently (overuse) or infrequently (underuse) compared to LOCNESS. Subsequently, a list of the most frequent bundles in TELC and LOCNESS was generated. To identify the cases of overuse and underuse related to the frequently occurring bundles across the two corpora, a Mann-Whitney U test was conducted as there were too many outlier values. To reduce the possibility of Type I errors, only the bundles with at least 20 occurrences were considered for the analysis.

3.3.2. Addressing research question 2

The second question aims to investigate the accuracy of 4-word PP-based bundles employed in by L1 Turkish L2 English students and how do they use regular expressions. The data analysis involved a two-step procedure. In the first step, a set of regular expression codes were added to the search query of Antconc. Then, the Regex function was selected to conduct the regular expression analysis on 86 reference bundles to find

deviations in the forms of target lexical bundles. The retrieved results were analyzed qualitatively to detect attempted bundles both in the form of errors (e.g., *in the other hand*) or different variations of the target lexical bundles (e.g., the use of *on one hand* instead of *on the one hand*). The errors were categorized into addition, omission, and misformation categories based on the framework of Dulay et al. (1982). The interrater reliability was assessed by using Cohen's Kappa.

The error analysis included all types of errors that are related to the target bundles form. Therefore, articles, prepositions, and errors related to the head nouns were all coded. All lexical bundle errors were classified independently and in cases of disagreement, the raters engaged in negotiations. Both errors and the other correct uses of the target bundles (e.g., *on one hand* instead of *on the one hand*) were taken as attempted bundles. As a last step, the attempted bundles were added to the overall lexical bundle counts.

3.3.3. Addressing research question 3

The third research question aims to explore how learners use the PP-based lexical bundles to gain insights from a developmental perspective. To investigate RQ3 the attempted bundles were added to the overall list of lexical bundles. More than half of the essays in TELC (n = 382) were scored in the ranges between 5 to 7. To increase the representativeness of proficiency levels, TELC was divided into low-scored (n = 314) and high-scored (n = 373) essays by running a K-means cluster analysis on writing performance scores. K-means cluster analysis ensured that essays scored between 5 to 7 were distributed evenly across the two groups. Overall RQ3 investigated two main aspects: (1) the extent to which the frequency of LBs based on prepositional phrases (PP) can predict writing quality, and (2) the differences in PP-based LB frequency between high-scored and low-scored essays.

To answer RQ3a, essays without bundles were not analyzed. To see if the frequency of PP-based bundles can predict total writing scores, low-scored essays, and high-scored essays, a linear regression test was used. Each group followed an 80/20 random stratified split. In more detail, the equation for the regression model was obtained from the training set while the test set was used to test the accuracy of the model. The resulting training set for each group is as follows: All scores (n = 198), high-scored essays

(*n* = 116), low-scored essays (*n* = 84). To meet the assumptions of the test, the data were tested for normal distribution among standardized residuals, linearity, multicollinearity, and homoscedasticity. There was no violation of assumptions regarding normal distribution among standardized residuals. The linearity test yielded significant results ($p < .05$). Also, the VIF score (1.00) followed the established standards of multicollinearity. Examination of scatterplots showed that independent variables were linearly related to the dependent variable, and any sets of data were not heteroscedastic. The analysis for all groups followed the same procedure. The accuracy of regression model was not tested for the groups with non-significant results. Following the analytical procedures, the test set was used to perform model accuracy indices. The equations (1-3) for performing the analytical procedures are given below:

$$Mean\ Absolute\ Error\ (MAE) = \frac{\Sigma\ (\mathrm{Y_i} - \widehat{\mathrm{Y}}_\mathrm{i})}{n} \quad (1)$$

$$Root\ Mean\ Squared\ Error\ (MSE) = \frac{\sqrt{(\Sigma(\mathrm{Y_i} - \widehat{\mathrm{Y}}_\mathrm{i})^2}}{n} \quad (2)$$

$$Mean\ Absolute\ Percentage\ Error\ (MAPE) = \frac{\left(\Sigma \frac{\mathrm{Y_i} - \widehat{\mathrm{Y}}_\mathrm{i}}{\mathrm{Y_i}}\right)}{n} * 100 \quad (3)$$

Note. $\mathrm{Y_i}$ = actual values, $\widehat{\mathrm{Y}}_\mathrm{i}$ = predicted values or fitted values, n = sample size

To address RQ3b, the data was quantitatively analyzed by comparing the frequency of tokens and types of lexical bundles across low-scored and high-scored essays. The essays without PP-based bundles were also included in this comparative analysis. The corpus was divided into low-scored (*n* = 314) and high-scored (*n* = 373) essays by running a K-means cluster analysis on writing performance scores. A Mann-Whitney U test was computed to provide statistical evidence for the differences between the two groups.

CHAPTER 4: RESULTS

This chapter is subdivided into three subsections, and the findings for each research question are presented separately. To address the first research question, the chapter will compare the use and frequency of PP-based reference bundles found in TELC and LOCNESS. To answer the second research question, it will then present findings on regular expressions and lexical bundle errors related to the use of PP-based bundles. For the last research question, the chapter investigates the relation between writing performance and PP-based lexical bundle frequency then it describes how PP-based bundles vary across low- and high-scored essays.

4.1. RQ 1: PP-based bundles used by TELC and LOCNESS

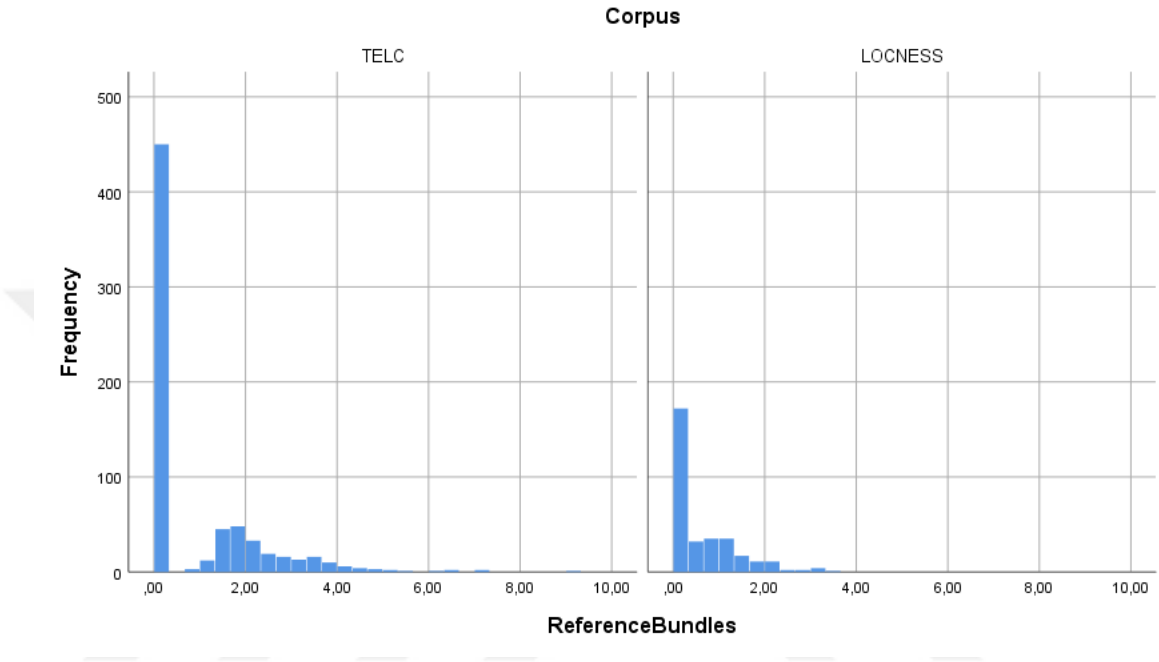
The first question aims to investigate the use of PP-based bundles in LOCNESS and TELC in terms of LB type and token distribution. Firstly, the chapter will present descriptive results on frequency counts. Then, the results on the inferential statistical tests will be presented. Finally, to analyze overuses and underuses of LBs, the section will compare frequency counts in two corpora and provide related inferential statistics.

4.1.1. Frequency of PP-based bundles in TELC and LOCNESS

TELC and LOCNESS were searched for 86 reference bundles. The frequency analysis revealed that more than half of the total essays in both corpora do not contain PP-based lexical bundles. However, the overall proportion of essays containing PP-based lexical bundles differs between the two corpora. Figure 4 is a cumulative graph of the number of lexical bundles found in TELC and LOCNESS. It reflects the frequency of bundles and the number of essays that do not contain reference bundles.

Figure 4

Distribution of Essays with PP-Based Lexical Bundles



As plotted in the y-axis of Figure 4, most students in TELC did not use the target structure. Specifically, 450 essays in TELC and 237 in LOCNESS do not contain PP-based LBs. The reason why two-thirds of the total essays in TELC do not contain target bundles might be related to the difference in text length, as they are shorter than those in LOCNESS. In TELC, the number of bundles was distributed evenly across the four essay tasks, with 60 for task 1, 70 for task 2, 67 for task 3, and 50 bundles for task 4.

To investigate the most frequently found bundles and shared bundles, raw frequencies related to reference bundle use was retrieved (Table 6). The table only lists bundles that occur in either corpus. Interestingly, the bundle in the pathogenesis of was not used by any learner groups, and the phrase in the treatment of only occurred once although both corpora included texts related to medical issues. According to Table 6, out of the total 20 out of 86 reference bundles were not used by either corpus. Notably, the bundles containing the noun *presence* and the bundle *in the pathogenesis of* were not used by any learner groups. Additionally, none of the learner groups preferred the bundles such as *in the next chapter*, *in the next section*, *in the study of,* and *in the present study*. These

bundles are discourse organizers that help readers to navigate through the text. Table 6 only lists the raw frequencies of reference bundles. The results regarding the overall tokens and types in TELC and LOCNESS and the related normalized frequencies are reported in Table 7.

Table 6

All Instances of PP-Based LBs in TELC and LOCNESS

Lexical Bundles	T*f*	L*f*	Lexical Bundles	T*f*	L*f*
on the other hand	91	30	**in the same way**	3	3
of the most important	58	9	from the fact that	2	0
at the same time	39	15	**of the use of**	2	2
as a result of	20	22	as a consequence of	2	0
to the fact that	17	18	**in addition to the**	2	1
of the fact that	12	2	**in such a way**	2	3
in the number of	8	6	**in the absence of**	2	0
at the end of	7	14	**in the process of**	2	1
from the point of	7	0	in this case the	1	0
at the beginning of	5	8	**for the development of**	1	3
in a variety of	4	1	**in the development of**	1	1
in a way that	4	8	**by the use of**	1	2
in the face of	4	2	**to the use of**	1	2
in the first place	4	3	about the nature of	1	0
on the one hand	4	2	**as a matter of**	1	1
in the use of	3	2	**as a means of**	1	3
for the purpose of	3	2	**in terms of the**	1	3
for the first time	3	3	**in the course of**	1	1
in the case of	3	26	**in the form of**	1	6
in the formation of	1	1	of some of the	0	2
on the basis of	1	1	over a period of	0	2

Lexical Bundles	T*f*	L*f*	Lexical Bundles	T*f*	L*f*
on the part of	0	6	to the start of	0	1
by the fact that	0	5	to the development of	0	1
in a number of	0	5	as part of the	0	1
in an attempt to	0	5	at the level of	0	1
as in the case	0	4	at the time of	0	1
by the end of	0	3	in the context of	0	1
in the area of	0	3	in the early stages	0	1
on the grounds that	0	3	in the sense that	0	1
to that of the	0	3	in the treatment of	0	1
to the extent that	0	3	similar to that of	0	1
as part of a	0	2	similar to those of	0	1
at the expense of	0	2	with respect to the	0	1

Note. Shared bundles are written in bold. Tf = Frequency in TELC, Lf = Frequency in LOCNESS

Table 7

Total PP-Based Bundle Types and Tokens in LOCNESS and TELC

	TELC (*n* = 687)		LOCNESS (*n* = 322)	
	R*f*	N*f*	R*f*	N*f*
LB Tokens	326	585.9	269	185.7
LB Types	39		62	

Note. Rf = Raw Frequency, Nf = Normalized Frequency per 500 words.

According to Table 7, TELC comprises 326 tokens and 39 types of bundles, whereas LOCNESS contains 267 tokens and 62 types of bundles. Therefore, while TELC contains more tokens of bundles, LOCNESS exhibited greater diversity of bundles. Table 6 illustrates that there are a total of 35 shared bundles. These shared bundles comprise 96% of the bundles in TELC (313 tokens) and 77% of the bundles in LOCNESS (207 tokens), suggesting that shared bundles rank high in frequency in both corpora. The proportional distribution of shared bundles could be influenced by the fact that LOCNESS utilized a greater number of bundle types. It is also worth noting that since the bundles are retrieved

based on a reference bundle list, the results regarding the proportional distribution of shared bundles might be inflated.

Furthermore, an analysis of the structural categories showed that a greater proportion of bundles in TELC were other PP fragments, accounting for almost 75% of the total bundles. On the other hand, structural categories of LBs were more evenly distributed in LOCNESS with PP with of-phrase fragments constituting 56% of the total bundles while other PP fragments covered the remaining 44%. According to the analysis of the structural distribution of bundles, it appears students in TELC write essays generally relied on adverbial bundles such as *on the other hand* and *at the same time*.

4.1.2. Differences in the use of PP-based LBs between TELC and LOCNESS

A Student's independent samples T-test was employed to further evaluate the difference between LOCNESS and TELC in terms of frequency of occurrence. The results were found to be statistically significant ($t(1003) = 3.08$, $p < .002$) with a small effect size ($d = 0.23$). The test revealed that essays in TELC ($M = 0.81$, $SD = 1.27$) made greater use of PP-based bundles than those in LOCNESS ($M = 0.57$, $SD = 0.76$). Since the most shared bundles are bundles that rank high in frequency, the list of 20 most frequent bundles were retrieved to further analyze the distribution of PP-based lexical bundles. Table 8 lists the 20 most frequent bundles in TELC and LOCNESS. In TELC, the three most frequently occurring lexical bundles are *on the other hand* (169.5 tokens), *of the most important* (107.3) and *at the same time* (68.2). These three bundles occurred 188 times in raw frequencies, comprising 57% of the total LB tokens in TELC. Moreover, the range of occurrences for *on the other hand* (84 in range), *of the most important* (53), and *at the same time* (33) demonstrates that the high frequency of these bundles is not due to idiosyncratic uses as they are dispersed across many texts.

Table 8

The Most Frequent Bundles in TELC and LOCNESS

TELC				LOCNESS			
Lexical Bundles	N*f*	R*f*	Range	Lexical Bundles	N*f*	R*f*	Range
on the other hand	169.5	91	84	on the other hand	21.6	30	28
of the most important	107.3	58	53	in the case of	16.9	26	22
at the same time	68.2	39	37	as a result of	16.9	22	19
as a result of	32.1	20	18	to the fact that	11.2	18	16
to the fact that	27.7	17	17	at the same time	10.2	15	15
of the fact that	17.7	12	9	at the end of	8.2	14	11
in the number of	16.1	8	7	of the most important	5.6	9	7
from the point of	11.3	7	7	in a way that	5.4	8	8
at the end of	10.7	7	6	in the number of	5.4	6	6
at the beginning of	8.5	5	5	in the form of	5.2	6	6
in the first place	8.4	4	4	on the part of	4.9	6	6
in a variety of	8.1	4	4	at the beginning of	3.2	8	7
on the one hand	7.7	4	3	as in the case	3	4	4
in the face of	7.2	4	4	as a means of	3	3	3
in the same way	7	3	3	by the fact that	2.8	5	5
in the use of	6.6	3	3	for the first time	2.7	3	3
in a way that	5.8	4	4	in an attempt to	2.6	5	4
for the purpose of	4.8	3	3	in a number of	2.5	5	4
for the first time	4.8	3	3	by the end of	2	3	3
in the case of	4.4	3	3	for the development of	1.6	3	2

Note. Rf = Raw frequency, Rf = Normalized frequency per 500 words.

In contrast to TELC, the top three bundles in LOCNESS only amount to 29% of all bundles found. The bundle *on the other hand* (21.6) is also the most frequently occurring bundle in LOCNESS. The second most frequent bundle in LOCNESS is *in the case of* (16.9) followed by *as a result of* (16.9). A close qualitative inquiry of concordance lines showed differences in the uses of the most frequently found bundles.

On the other hand

The most frequent bundle in both groups is *on the other hand*; it can also be considered as the bundle with widest dispersion. It is a linking adverbial with organizational function that connects priori and following discourse. In TELC, out of 91 *on the other hand* bundles 77 bundles (84% of the total occurrences) were found to be used as a clause-initial bundle (1a). In LOCNESS 63% (19 tokens out of 30 bundles) of these bundles were found to be used as clause-initials. Moreover, in TELC, the text that contained *on the other hand* mostly included its collocate *on the one hand* (1c).

(1a) It has numerous contributions to humanity. *On the other hand*, it has some shortcomings. (TELC)

(1b) Unorganized crime does not pay, at least not very well. Organized crime, *on the other hand*, pays off quite handsomely. (LOCNESS)

(1c) I think, *on the one hand*, it makes our life easier, *on the other hand*, it makes us deserialize from many things. (TELC)

of the most important

In both groups the bundle *of the most important* was commonly used in the form of noun phrase + post modifier fragments (i.e., *one of the most important*). When used as a 5-word noun phrase-based bundle it emphasizes a point and remarks on the significance of the message (see example 2a). Therefore, this structure was mostly used as a stance expression.

(2a) Entertainment is *one of the most important* causes of isolation. (TELC)

Surprisingly, this bundle appeared 33 times in task 3 essays, representing more than half of its overall token counts (58 tokens) in TELC. It was also observed that in TELC, *of the most* structure was used with other nouns such as *essential* (9 tokens) as given in example

2b, and *basic* (5 tokens) as illustrated in example 2c. In one case, *of the most important* was used by taking an additional preceding adjective such as *crucial* (2d). Such uses are considered as erroneous use of the target structure and will be discussed in the regular expression analysis section.

(2b) That is why empathy is one *of the most essential* social/personal skills that one needs. (TELC)

(2c) The Internet is considered *one of the most basic* needs of our lives. (TELC)

(2d) Empathy is one *of the most crucial important* skills in human life. It gives people a viewpoint from different angles. (TELC)

At the same time

In TELC the third most frequently used bundle is *at the same time*. It can suggest a temporal relationship (3a) and signal contrast (3b). The bundle *at the same time* was used 22 times to suggest temporal relationship (i.e., as a time adverbial) and 17 times to signal contrast out of 39 instances. In LOCNESS it was used 6 times to indicate temporal relationships and 9 times to express contrast. Furthermore, the bundle was used as a sentence-initial bundle 16 times in TELC, compared to only 5 times in LOCNESS.

(3a) They can listen to the teacher and the music *at the same time* if it helps develop their learning. (TELC)

(3b) Internet is the invention of all time. *At the same time*, it is the worst. (TELC)

In the case of

Compared to TELC, students in LOCNESS used the bundle *in the case of* more frequently (3 tokens versus 26 tokens) and over a wider range of texts. This finding suggests an unfamiliarity with this bundle in TELC. *In the case of* is a bundle that helps maintaining textual coherence as they help framing an argument or a point. It can be confused by the *in case of*. TELC includes an example (example 9) that demonstrates the incorrect use.

(4a) The online education system is a type of system used *in the case of* an unusual event like a pandemic outbreak. (TELC)

As a result of

Another bundle with high frequency in both corpora is *as a result of*. This bundle is the fourth most common in TELC (32.1) and the third most common in LOCNESS (16.9) in terms of frequency rank. In both groups, the bundle *as a result of* was often followed by definite articles or demonstratives, marking a connection to the contexts preceding them (5b). It denotes a consequence or effect of an action or event (5a).

(5a) This discussion arose after lots of boxers started to get serious injuries and brain damages *as a result of* boxing. (LOCNESS)

(5b) To begin with, in the 21th century it is hard to find a way to have fun in a cheaper way. *As a result of* this, people prefer going online rather than going outside because it is more affordable. (TELC)

As a result, which can be considered as the third-word form of as a result of was found to be used 81 times in TELC and 14 times in LOCNESS. Therefore, learners show a great tendency to use this bundle as a third word bundle instead of using the bundle *as a result of,* which contains two prepositions.

To the fact that

The lexical bundle *to the fact that* occurred 17 times in TELC and 18 times in LOCNESS, emphasizing its significance for closer inspection. In addition to the similarities of frequency counts, in both corpora to the fact that was often accompanied by "*due*", forming into 5-word bundle *due to the fact that.* In cases like in extract 6a, the word "*because*" can be substituted for this phrase to avoid verbosity, especially for shorter essays.

(6a) People start to maintain different lifestyles *due to the fact that* the internet changes their goals, work environment, economic or family lives. (TELC)

The differences in the normalized frequencies of bundles between two corpora point to the overuse and underuse of these features. The Mann-Whitney U test showed significant differences in the two groups' use of 6 bundles while the bundle *at the same time* yielded non-significant results (see Table 9). Therefore, out of 86 reference bundles, only 7 bundles met the assumption of sample size.

Table 9

Cases of Overuse and Underuse

Overuse	on the other hand, of the most important,
Underuse	in the case of, as a result of, to the fact that, at the end of
Non-significant	at the same time

Note. For overuse and underuse only the level of p < .05 was considered.

In LOCNESS there was a greater use of 4 bundles, including *as a result of, to the fact that, in the case of, at the end of*. In contrast, only two bundles were more frequently used by TELC: *on the other hand*, and *of the most important*. Interestingly, while the raw frequencies of *as a result of, to the fact that* varied little across both corpora (as shown in Table 6), Table 10 reveals that LOCNESS made a greater use of these bundles than TELC. Overall, a small effect size ($d < 0.20$) is observed for each bundle listed for overuse and underuse.

Table 10

Statistically Significant Differences in Lexical Bundle Use

Bundles	*U*	*d*	*p*
on the other hand	115549.000	0.06	.036
of the most important	116908.500	0.11	< .001
as a result of	106956.000	-0.08	.008
to the fact that	107968.500	-0.06	.047
in the case of	103562.000	-0.19	< .001
at the end of	107825.500	-0.09	.004

Note. Only the cases at p < .05 were considered.

A total of 79 bundles occurred rather infrequently in both TELC and LOCNESS, with 26 of them appearing only once or twice. Table 11 lists the bundles that were not included in the Mann-Whitney U analysis as they violated the assumption of sample size.

Table 11

PP-Based Lexical Bundles Excluded from The Mann-Whitney U Analysis

Bundles that occur infrequently	about the nature of, as a function of, as part of the, as a consequence of, as a matter of , as a means of, as part of a, at the time of, at the beginning of, at the level of, at the expense of, to the start of, by the end of, by the presence of, by the end of, for the development of, for the purpose of, for the purposes of, from the point of, in the case of, in the absence of, in the form of, in the presence of, in a number of, in terms of the, in the context of, in the course of, in the development of, in the number of, in the process of, in a variety of, in the area of, in the direction of, in the face of, in the formation of, in the pathogenesis of, in the study of, in the treatment of, in the use of, in view of the, of a number of, of some of the, of the effects of, of the nature of, of the use of, on the basis of, on the surface of, over a period of, to the development of, to that of the, to the presence of, to the use of, with the exception of, as in the case, between the two groups, by the fact that, for the first time, from the fact that, in such a way, in the same way, in the present study, in a way that, in addition to the, in an attempt to, in contrast to the, in relation to the, in the early stages, in the first place, in the next chapter, in the next section, in the sense that, in this case the, of the fact that, on the one hand, on the grounds that, similar to that of, similar to those of, to the extent that, with respect to the

4.2. RQ2: Accuracy of lexical Bundles

The second question analyzes how learners in TELC used regular expressions to identify potential challenges learners face in the formation of PP-based lexical bundles.

4.2.1. Regular expression analysis

The first stage of the analysis was quantitative and involved the identification of regular expressions using Antconc's Regex search query. The regular expression search parameters retrieved 1209 tokens on 49 types related to the use of reference lexical bundles. The frequencies listed are the raw frequency of tokens. A notable difference can be seen in Table 12 in that TELC has considerably more regular expression hits (49 types,

1209 tokens) compared to the initial search for lexical bundles (39 types, 326 tokens; Table 7). Interestingly, despite being among the most frequently occurring bundles, *on the other hand* and *at the same time* were among the least frequently found regular expressions. The regular expression search retrieved the deviations from the reference bundles, which are three-word forms of the reference bundles, reference bundles in different forms, and learner errors. Most of the regex patterns that occurred over 50 times were three-word versions of some of the reference bundles.

Table 12

Regular Expressions in TELC

Lexical Bundles	*f*	Lexical Bundles	*f*	Lexical Bundles	*f*
as part of (the or a), on the part of	190	in the case of	10	with respect to the	2
in terms of the	104	(at or by) the end of	8	in view of the	2
(in, of, to, or by) the use of	101	as a consequence of	8	in the treatment of	1
as a result of	87	in the absence of	6	in the study of	1
in (a or the) number of, of a number of	86	in the form of	6	on the surface of	1
in addition to the	76	in the course of	6	as a function of	0
of the most important	72	in a variety of	6	at the expense of	0
of some of the	71	as a matter of	5	(by, in, to) the presence of	0
(by, to, from, or of) the fact that	55	in the face of	4	as in the case	0
of the effects of	34	about the nature of, of the nature of	4	in the formation of	0
(for, to, or in) the development of	32	over a period of	4	in the pathogenesis of	0
in the area of	31	in an attempt to	4	in the next chapter, in the next section	0
at the time of	25	in contrast to the	4	similar to that of, similar to those of	0
to that of the	23	in the same way	4	in relation to the	0
from the point of	17	in a way that	4	in the sense that	0
on the (one or other) hand	16	on the basis of	4	to the extent that	0
for the purpose(s) of	16	at the same time	3	in the present study	0
at the level of	16	to the start of	2	in such a way	0
in the process of	14	in the context of	2	on the grounds that	0
in this case the	12	in the direction of	2	in the early stages	0
as a means of	10	between the two groups	2	in the first place	0
at the beginning of	10	for the first time	2		

These three-word forms of the reference bundles include *part of the*, *in terms of*, *the use of*, *as a result*, *the most important*, *the number of*, *in terms of*, *the fact that*, and *in addition to*. The sentence in (7a) contains two of these three-word forms of the target bundles. In such examples, the bundles are used correctly but not in the form of the reference bundles searched.

(7a) Thus, *the use of* books and notebooks is reduced. *As a result,* many people prefer hybrid education.

There were also the uses of reference bundles but with a plural form the noun that it contains. As illustrated in the sentences below (7b and 7c), students used "*facts*" and "*times*" instead of the singular forms of these head nouns.

(7b) They could put limits *at the times of* using cars or they could give as a gift bicycles to people to use them rather than their cars. (at the time of)

(7c) It helps an individual to manage his feelings even *at times of* great stress. (at the time of)

In extract 7b the correct use of the bundle (i.e., *at the time of*) might not make the sentence semantically meaningful, so it cannot be considered as an attempted bundle. Example 6c, on the other hand, is an attempted bundle since it is used correctly in context, but not in the form of the target reference bundle. In total, the regular expression analysis found 8 attempted bundles that are used correctly, but do not match the targeted reference bundles in structural form. The following sentences (7d-7g) show the instances of these bundles.

(7d) *On one hand*, cars cause air pollution with their poisonous gas. (on the one hand)

(7e) *On one hand*, from an artistic perspective, as simplest when people read a book or a poem, they try to understand their meanings, the feelings of the story. (on the one hand)

(7f) *Due to fact* that the corona virüs has swept the whole World, all countries first switched to online education, then turned into hybrid system. (to the fact that)

(7g) They can pay their bills, do shopping, play an online game with their friends *at same time* but different countries or chat and arrange a meeting with just a computer and Wi-Fi connection. (at same time)

4.2.2. Incorrect uses of PP-based bundles

To further identify the attempted bundles in learner essays, erroneous uses of lexical bundles were identified. Table 13 lists the frequencies related to the three categories of errors in the use of PP-based lexical bundles. Interrater reliability for error identification and coding was high ($\kappa = 0.89$, $p < .0001$).

Table 13

Three Types of Errors with PP-Based Lexical Bundles

Lexical Bundles	Omission	Addition	Misformation	Total
on the other hand	-	-	3	3
on the one hand	-	-	2	2
to the development of	1	-	1	2
in the development of	1	-	-	1
for the development of	1	-	-	1
at the same time	1	-	-	1
in the same way	1	-	-	1
as a matter of	-	-	1	1
at the beginning of	1	-	-	1
at the level of	-	-	1	1
in the form of	1	-	-	1
in addition to the	1	-	-	1
of the fact that	1	-	-	1
between the two groups	1	-	-	1
TOTAL	10	-	8	18

The table above shows that there are 18 errors within 12 lexical bundle types in Turkish students' essays. According to the results of the analysis, none of the errors were in the category of addition. Surprisingly, the inaccurate uses of bundles were predominantly associated with bundles that have idiomatic meanings (e.g., *on the one hand* and *on the other hand*). The errors within the bundles, *on the one hand*, *on the other hand*, *at the same time*, *in the same way*, accounted for 38,8% of all errors found in the analysis. Overall, the most frequent errors were associated with the misformed use of *on the other hand*, *on the one hand* and *in the development of*. Examples 8a and 8b illustrate the erroneous uses of *on the one hand*.

(8a) *In one hand* cars it is very important way fot traffic yes this is true and the most of people say this because they can't go to work or school easily.

(8b) *On a hand*, excessive use of the Internet in the social sense, getting away from the outside/real world so people are being more individual, antisocial, internet addicted personality.

There were also errors related to *on the other hand*, which is a collocate of *on the one hand*. In example 7c, the student confused the appropriate preposition while in examples 7d and 7e students choose other head nouns rather than *hand*. However, as can be seen from examples 7d and 7e both cases are used as a clause initial bundle to present a contradictory opinion. Therefore, these examples can be considered as misformed use of the bundle *on the other hand*.

(7c) *in the other hand* cars should be not banned yes i think also this because cars it is very easy way for transport şn tecnology world for example if they have hastily and you have car this is very easi for you.

(7d) *On the other way* one reason for not banning is that modern times people they have very interest at cars and if they banned they dont go here to spend time.

(7e) *On the other side*, crowded cities for example; İstanbul,New York have quite a lot of opportunities in terms of business.

Another commonly confused bundle in student essays is *to the development of*. Students omitted the definite article in example 7f and omitted the preposition in example 7g.

(7f) Educational models are always chancing according *to development of* technology and social problems.

(7g) This may lead *to the development* not only individuals but also all the society.

Since the study is concerned with the errors in PP-based bundles, the total number of errors includes both preposition and article errors found in the target bundles. In total, 9 of the errors listed in the table were related to the erroneous uses of prepositions, while 7 were due to the incorrect use of articles, and the remaining 2 errors were related to the misformation of head nouns. Therefore, the accuracy rate of bundles in learner essays is 94.9%. As summarized in Table 14, the regular expression analysis identified a total of 26 attempted bundles in 4 lexical bundle types.

Table 14

The Frequency of Types and Tokens with Attempted Bundles

	TELC	LOCNESS	TELC with attempted bundles
LB Tokens	326	269	352
LB Types	39	62	43

4.3. The Relation of PP-based Bundle Use to Writing Quality

The third question is two-fold: RQ3a investigates the effects of 4-word PP-based bundle frequencies on writing quality, while RQ3b analyzes how frequency counts differ across high- and low-scored essays.

4.3.1. The effects of PP-based bundle frequency on writing performance

The RQ3a investigated whether PP-based bundle frequency can predict writing scores in the whole corpus, in high scored essays, or in low-scored essays. Only the essays that contain 4-word PP-based bundles (including attempted bundles) were considered for the analysis, so the essays with no bundles were excluded. For each group, the data was split based on 80/20 ratio following a stratified random sampling. As shown in Table 15, the initial analysis on PP-based bundle frequencies and total essay scores was statistically significant only for the low-scored essays. Therefore, accuracy of the regression model was only calculated for low-scored essays by dividing the data into training (n = 84) and test sets (n = 21) based on 80/20 ratio.

Table 15

Regression Model Summary for All Groups

Group	Variable	*SS*	d*f*	*MS*	*F*	*p*
All_Scores	Regression	1.660017	1	1.660017	.907	.342
	Residual	358.699515	196	1.830100		
	Total	360.359533	197			
High_Scored	Regression	1.156462	1	1.156462	.521	.215
	Residual	84.977159	114	.745414		
	Total	86.133621	115			
Low_Scored	Regression	2.364072	1	2.364072	4.441	**.038**
	Residual	43.646345	82	.532273		
	Total	46.010417	83			

Note. Sample sizes for three groups are as follows: All_scores (n = 197), High_scored (n = 116), Low_Scored (n = 84).

Table 16

The Results of Coefficient Analysis for Low-Scored Essays

Variable	*B*	*SE*	β	*t*	*p*
(Constant)	5.533	.191		28.963	< **.00001**
Low_Scored	-.138	.065	-.227	-2.107	**.038**

Coefficient analysis on training set yielded a significant model (R= .24, R^2 = .051, $F(1,83)$ = 4.441, p = .038), explaining 5% of the variance in the evaluation of 84 low-scored essays in TELC (see Table 16). Also, there is a negative correlation (r = - .227 p = .019) between writing performance and PP-based lexical bundle frequency in low-scored essays. Accordingly, the regression equation for low-scored essays was built as follows:

$$\text{Low\_scored} = 5 \cdot 533 + (-.138 * \text{independent\_variable})$$

The independent variable in this case is the frequency of PP-based bundles found in the training set of low-scored essays. A subsequent analysis on test set for accuracy (see Table 17), returned Root Mean Squared Error (RMSE) of 0.8720, indicating 91.27% accuracy in the prediction of essay scores.

Table 17

Summary of Regression Model Accuracy (n_{Low} = 21)

Accuracy Measures	Error	Accuracy	Accuracy Percentage
Mean Absolute Error	0.167857	9.832143	98.32143
Root Mean Squared Error	0.872044	9.127956	91.27956
Mean Absolute Percentage Error	14.74651	85.25349	85.25349

4.3.2. Frequency of PP-based bundles used by high- and low-scored essays

The descriptive findings of RQ3b revealed differences in the frequency counts of LBs in high- and low-scored essays. Table 18 summarizes the distribution of PP-based lexical bundle types and tokens in high-scored and low-scored essays.

Table 18

Bundle Types and Tokens in High- and Low-Scored Essays

	High-Scored (*n* = 373)		Low-Scored (*n* = 314)	
	R*f*	N*f*	R*f*	N*f*
LB Tokens	214	354.6	136	272.1
LB Types	39		30	

Note. Rf = Raw Frequency, Nf = Normalized Frequency.

A Mann-Whitney U test was employed to compare the mean values between the two groups. The Mann-Whitney U test revealed no statistically significant differences in the normalized frequency counts across the two groups (p = .46). There were also no significant results on the distribution of lexical bundle types (p = .18). To further investigate the difference in frequency counts and bundle types across two groups, the most frequent bundles from each group was investigated. Highly frequent bundles such

as *on the other hand,* showed similar distribution in each group with no statistical differences worthy of note. In terms of the structural subcategories of PP-based bundles, low-scored essays contained slightly fewer PP with *-of* bundles compared to high-scored essays. In detail, 20% of the bundles in low-scored essays and 30% of the total bundles in high-scored essays contained PP-based bundles ending with *-of*. Consequently, it was found that there is no evidence of overuse or underuse.

CHAPTER 5: DISCUSSION

5.1. Frequency analysis of PP-based bundles

To determine the frequency of PP-based lexical bundle types and tokens in the essays of TELC (n = 687) and LOCNESS (n = 322), both corpora were searched for 86 reference bundles. It was found that approximately one-third of the essays in TELC essays and half of the essays in LOCNESS contained PP-based lexical bundles. Despite the rate of essays without the use of bundles, the initial frequency analysis retrieved 326 tokens (normalized frequency of 585.9) and 39 types of LBs in TELC and 269 (normalized frequency of 185,7) tokens and 62 types of LBs in LOCNESS. Based on the frequency-driven analysis, there seems to be a pattern of greater use of LB tokens by TELC, while LOCNESS utilized a greater variety of bundle types. An independent samples T-test provided statistical evidence for the differences in the use of LBs between TELC and LOCNESS ($p < .002$, $d = 0.22$), suggesting Turkish EFL learners used PP-based bundles slightly more frequently. A further statistical analysis using Mann-Whitney U showed that particularly the bundles *on the other hand* and *of the most important* were overused in the essays of TELC, while the bundles *at the end of, in the case of,* as *a result of,* and *to the fact that* were underused. In addition, the two overused bundles account for approximately 50% of all bundles found in TELC, supporting the idea that Turkish learners are using these bundles repetitively. The overreliance on certain bundles was also reflected in the findings related to the shared bundles, as almost all bundles found in TELC were shared by those in LOCNESS. Overall, the Turkish EFL writers demonstrated more uses of lexical bundles in number but lacked variety in their uses compared to LOCNESS.

Previous studies partly agree with the results of this study. The previous body of studies has shown that non-native writers use fewer tokens of lexical bundles (Lee et al., 2020, Shin, 2018, Shin et al., 2018). Counterintuitive findings could be attributed to the overuse of certain bundles such as *on the other hand* and *of the most important*, as they inflate the number of bundle counts. In parallel with the aforementioned studies, this study found less diverse use of PP-based LBs in Turkish L2 learners compared to their native

counterparts. The overuse of two bundles (i.e., *on the other hand*, and *of the most important*) accounts for over half of the total LB count found in TELC, indicating that learners use these bundles redundantly. In a similar vein, Lee et al. (2020) reported that L2 learners tend to rely on a small repertoire of lexical bundles (e.g., *of the most important*) and use them repetitively while underusing others, such as *in the case of*. In another study, Yoon and Choi (2015) compared the lexical bundles in Korean students' writing to LOCNESS by analyzing lexical bundle type-token ratios and found that essays in LOCNESS showed a greater variety of PP-based bundles. Studies exploring the lexical bundle use among professional NNS writers and professional NS writers (e.g., Wei & Lei, 2011) also yielded comparable results regarding the overuse of a restricted number of bundles. Therefore, this pattern might not be unique to Turkish EFL learners or novice academic writers but rather be explained by the nature of non-native phraseology.

As pointed out by Hasselgård (2019), language learners commonly rely on familiar multi-word units that they feel comfortable using, which she refers to as "*phraseological teddy bears*". Hasselgård found that even advanced L2 learners can overuse high-frequency basic lexical bundles and ignore alternatives rather than risk making mistakes. In TELC, the bundle *on the other hand* was found to be overused and was primarily utilized as a clause-initial bundle, creating a compelling argument for its classification as a phraseological teddy bear. In addition, bundles such as *on the other hand* and *at the same time* are psycholinguistically salient bundles as they are complete structures that are glued together. Therefore, learners are more likely to acquire them early in their language development as they entail a clear meaning (see Conklin & Schmitt, 2008). The prevalence of phraseological teddy bears and the limited variety of bundle types partly explain why the current study found more tokens of bundles in TELC compared to LOCNESS. However, it is also worth noting that the overuse of the bundle *of the most important* might be related to students' prompt dependency. Because this bundle predominantly occurred in task 3 essays of TELC, where the prompt includes the sentence: "Empathy is considered to be one of the essential personal/social skills in the 21st century.", which can potentially account for its overuse.

This study also found underused bundles such as, *at the end of, in the case of, as a result of,* and *to the fact that.* As can be seen, compared to highly frequent bundles such

as *on the other hand* or *at the same time*, these bundles are neither psycholinguistically salient nor complete grammatical structures. The underuse of *in the case of* might be related to Turkish EFL learners' register awareness as these bundles are highly frequent in academic registers. According to Biber et al. (1999), *in the case of* is one of the top two most common bundles used in academic prose. Fewer instances of *in the case of* might also be related to the complexity in the production of these bundles as it contains two prepositions. This is also the case for *as a result of* and *at the end of*. In fact, there were considerably more instances of *as a result* in TELC (81 tokens) than in LOCNESS (14 tokens), supporting the idea that learners struggle to produce longer bundles. Except for the bundle *to the fact that*, the other underused bundles were framing devices that serve a pragmatic function to identify specific attributes. According to Yoon and Choi (2015), although students attempt to create logical relationships by using transitions, they fail to make logical inferences as they rarely employ framing bundles. The same conclusion holds for the present study. Lastly, the underuse of *to the fact that* might be due to the difference in the textual composition of the two corpora compared, as in TELC the essays are relatively shorter. Since both corpora used the bundle *of the fact that* as *due to the fact that,* it is not rare for shorter essays to contain fewer instances of this bundle.

It is important to note that the analysis excluded 79 bundles from the Mann-Whitney U analysis as they violated the assumption set for minimum sample size. These bundles occurred too infrequently to produce reliable results. Notably, 26 LBs appeared only once or twice in either corpus, while 20 of them were not produced by any of the learner groups. Notably, the bundle with the head noun *presence*, and the bundles such as *in the next chapter* were not used by any learner groups. According to Hyland (2008), the LBs with the head noun "*presence*" were highly frequent in texts on hard science topics (e.g., biology and engineering). Since the essays considered in this study are not on hard science topics, but rather covered more general subjects, it is natural for learners not to use these types of lexical bundles. Also, discourse organizers such as, *in the next chapter* and *in the present study* were not present in either corpus. Since students in LOCNESS and TELC only represent novice academic writers and are not at the same level of academic maturity as professional academics, they might not be familiar with these structures. However, these bundles help navigate the reader through the text

(Hyland, 2008) and are important constructs for framing and an argument. Therefore, focused instruction on these types of bundles might help students to be more familiarized with these bundles.

The subcategories of PP-based bundles varied between the two corpora. The bundles such as *on the other hand*, *at the same time* are categorized into other PP-based fragments in terms of their structure (Biber et al., 1999). Most PP-based bundles used were other PP fragments in TELC, whereas LOCNESS contained PP-based bundles with -of fragments and other PP fragments roughly evenly. When the structural categories of bundles are considered, PPs with -of fragments are highly prevalent and cover the majority of PP-based bundles. Chen and Baker (2016) suggested that it is not typical of academic writing to include high occurrences of adverbial bundles while underusing PPs with -of structure. In addition, Vo (2019) asserts that as learners advance in proficiency, they start to use lexical bundles beyond their adverbial meanings. Therefore, the findings regarding the distribution of structural subcategories of PP-based bundles further support the claims of underuse and overuse of bundles.

Overall, the findings of the present study contribute to Bestgen and Granger's statement (2014, p. 29), "L2 writers rely on a more limited repertoire of lexical bundles than native writers; they overuse the bundles they are familiar with, often calqued on similar sequences in their L1". In this study, the tendency to use lexical bundles repetitively might be related to their limited repertoire of lexical bundles and their dependency on prompt-based bundles. Some earlier studies, Staples et al. (2013) and a review of formulaic language by Granger and Paquot (2012) had a similar conclusion. As discussed in the previous literature, the overuse of a limited range of bundles adds redundancy and verbosity to their texts. This study is not without limitations: while all the essays in TELC were timed essays, LOCNESS contained both timed and untimed essays. Task setting might have influenced students' production of linking adverbials and connectives as time constraints might pressure learners to produce more of these structures (see Paquot, 2010). However, the appropriate use of lexical bundles is an integral part of academic writing, and instructors should emphasize the production of the appropriate use of these features. Also, there is a need for reevaluation of the content of the current textbooks in practice. As highlighted by Gedik and Kolsal (2022), the official

public secondary school teaching materials are not rich enough in syntactic and lexical complexity to prepare students for tertiary education. It is crucial to use authentic teaching material, ideally corpus-informed resources to raise awareness about alternative expressions for overused bundles and encourage the usage of underused bundles.

5.2. Erroneous uses of PP-based bundles

The second research question investigated the learners' erroneous uses of PP-based lexical bundles in context. The current study spotted instances of attempted, but deviated forms of LBs using regular expression codes. Such attempted bundles are ignored by the corpus- and frequency-driven methods, which are commonly employed to identify lexical bundles. Regular expression analysis retrieved 1209 results on 49 categories of PP-based lexical bundles. The results included uses of PP-based bundles deviating from the target form (e.g., *with the development of*) and 3-word LB forms of the 4-word reference bundles (e.g., *in addition to*). Out of 1209 regular expressions, there were only 26 instances of potential reference bundles, in which learners attempted to use the target PP-based lexical bundle forms. Attempted bundles included correct uses of reference bundles but in alternative forms (e.g., *on one hand*) and erroneous uses of reference PP-based bundles (e.g., *in the other hand*). The discrepancy between the number of attempted bundles and the regular expressions found shows that learners do not use the regular expressions of PP-based bundles in the target reference bundle forms. After the extraction of regular expressions, an error analysis was conducted on erroneous uses of lexical bundles (18 tokens) by two coders ($\kappa = 0.89$, $p < .0001$). The accuracy rate of bundles in learner essays was 94.9%, indicating that learners mostly used the reference bundles correctly.

The results showed that even though the Turkish language does not have preposition or article systems like English (see Underhill, 1976), the error rate in TELC was surprisingly low. While most common errors were related to the omission of articles, students did not commit any article addition errors. From a developmental perspective, addition errors are argued to occur at later stages of L2 acquisition (Dulay et al., 1982). However, according to Shin (2018), L2 English learners with article-less first language

predominantly use lexical bundles with bare head nouns without using preceding or following articles. Therefore, the lack of addition errors and prevalence of omission errors might be related to the nature of learners' L1 background. Consistent with this, the omission of articles was the most frequent error in the present study. In an analysis of Turkish-speaking EFL students' use of 4-word bundles, Uzun (2018) also reported omission to be the most frequent definite article error while learners did not commit any addition errors. Nevertheless, it is hard to firmly conclude whether learners largely used articles correctly or not, but consistent findings on the omission of definite articles add to the possibility of Turkish learners facing difficulties using these function words.

In terms of preposition errors, the findings of this study echo those of Lee et al. (2020), which also found Korean students to have a high rate of accuracy in their use of LBs. However, as mentioned earlier, the learners in TELC tended to use a limited repertoire of bundles and exhibited a lack of diversity. Therefore, learners might have employed an avoidance strategy to make fewer mistakes in their writing, as they avoid using the lexical bundles that they were not feeling comfortable using. Students in TELC particularly overused the bundles *on the other hand*. According to Granger (2017), using contrastive interlanguage analysis, it is possible to detect learner errors such as *in the other side* and *in the other hand,* which corresponds to the bundle *on the other hand.* Although bundles *on the other hand* and *on the one hand* entail a relatively clear meaning, Turkish learners frequently made errors using them. In fact, errors related to *on the other hand* and its collocate *on the one hand* accounted for one-quarter of all errors found. Among the errors related to *on the other hand,* there were uses such as *on the other side* or *on the other way,* pointing to the interference of L1. Turkish functional equivalent form of *on the other hand* is "*diğer taraftan*" and "*öte yandan*", which can translate into on the other side or on the other way. Though it is semantically transparent (Yorio, 1989), the bundle *on the other hand,* is not only overused but also misused. In L1 acquisition, conventionalized routines are argued to be acquired early, however, as shown in the findings adult learners struggle to make extensive use of the most commonly found bundles. From a usage-based perspective, the findings of this study contribute to the argument that adult L2 learners have difficulties in the use of highly conventionalized bundles (see Yorio, 1989).

The appropriate use of *on the other hand* in academic writing is vital as it is one of the most frequently used bundles in academic discourse. In addition, it functions to contrast two ideas and establish semantic links between the priori and coming discourse, further emphasizing its role in argumentative texts. Although cohesive devices can help foster cohesion and establish the connection between two units of discourse, fewer cohesive devices are associated with higher proficiency (Crossley & McNamara, 2012). Crossley and McNamara (2012) assert that learners at higher proficiency levels assume their readers to be higher-knowledge readers and do not need extensive connective devices to navigate through their readings. High frequency of sentence connectors such as *at the same time* and *on the other hand* provides a rationale for investigating the effect of semantic and formal misuse of sentence connectors and the role of cohesion on writing quality. Regular expression analysis can also provide insights into the use of sentence connectors as a similar code set can be generated for both phrase-level and clause level connectors found in Biber et al. (1999).

Aside from the errors found, there were attempted bundle uses that can be considered correct but go undetected in traditional frequency-based corpus investigations. When such uses are detected, the overall lexical bundle count can be refined to produce results that differ from those of the initial frequency-based analysis (Shin, 2018). For example, a student used *at times of* instead of *at the time of*. Although used correctly, it went undetected in the initial analysis. Thus, regular expression analysis can produce more accurate results on the overuse and underuse of bundles. This study only found 26 attempted bundles, but in a larger corpus, it is possible to quantify more attempted bundles that go undetected. Regular expression analysis also sheds light on the plural markings and three-word forms of reference bundles. Therefore, combined with the frequency-based analysis, regular expression analysis not only gives information about the adjoining elements of the LBs (e.g., the word *due* was commonly used with *to the fact that*), but also its use of shorter or colloquial forms (e.g., *on one hand* instead of *on the one hand*). Although the corpus used in this study is small, more studies using regular expression analysis can shed light on the frequently used three-word versions of 4-word bundles by language learners. These uses can be included as tips for the learners in coursebooks to raise awareness of alternative uses of bundles. Also, coursebooks can contain instances

of typical learner errors. For example, as Granger (2015b) demonstrated, the Italian version of *The English in Mind series* includes "Get it Right!" tip boxes that contain examples of common errors made by Italian learners.

In summary, regular expression analysis revealed that the error rate in reference bundles was quite low, indicating that students mostly used PP-based bundles with high accuracy. The error analysis showed that Turkish students tend to omit definite articles and misuse or omit prepositions. It is also worthy of note that there were L1-influenced errors related to the use of *on the other hand*, which is the most frequent bundle found in TELC. These results were mostly in line with the previous studies (e.g., Lee et al., 2020; Shin et al., 2018; Uzun, 2018). The size of the corpus used might be considered a limitation, as a larger corpus might have higher statistical power. Another limitation to be acknowledged is that the author did not investigate the semantic properties of the bundles used. An investigation of semantic misuse (e.g., sense relation errors) of bundles might produce a better understanding of the cause of errors. The occurrence of these errors has some pedagogical implications. Since the usage-based perspective posits that grammar is the cognitive organization of one's experience with language, the teaching of structures such as prepositions and articles within necessary contexts might help learners to interpret their function and use (see Gilquin, 2022). One way to teach language structures in context is by corpus tools or corpora. The book by Pinto et al. (2023) offers detailed guidance on how to use corpus tools in teaching practice with relevant lesson plan examples. Using these examples, innovative classroom activities can be designed and implemented to help students notice their errors.

5.3. The use of PP-based bundles and writing quality

5.3.1. The relationship between PP-based LBs and L2 writing quality

Durrant (2019) and Kim and Kessler (2022) claimed that the relationship between the use of FL and human holistic scores remains unclear and further studies are needed. To address this issue, RQ3a investigated the relation between PP-based bundle frequency and human holistic scores. According to the analysis, there was no significant relation between PP-based bundle tokens found in all essays and writing quality. In line with the

present study, previous studies (e.g., Kılıç, 2015; Kim & Kessler, 2022; Torlak, 2020) have also reported no statistically significant relation between frequency counts of formulaic sequences and overall writing performance. In light of these results, the frequency of PP-based bundles does not appear to distinguish low-scored essays from those of high-scored ones. In Kim and Kessler (2022), PP-based bundles distinct to the high proficiency learners were 3-word bundles. Their study demonstrated that an analysis of 3-word bundles might produce different results. It is also worthy of note that Kılıç (2015) investigated the relationship between frequency of formulaic sequences and overall writing scores without dividing the corpus into categories such as high-scored or low-scored essays. Staples et al. (2013) argued that formulaic language is an essential device for low-proficiency learners. In addition, Paquot (2018) demonstrated that phraseological competence could effectively distinguish between higher intermediate and advanced learners. To contribute to these arguments, as a part of the RQ3a, the present study analyzed the predictive power of PP-based bundle frequency in low-scored and high-scored essays.

To further investigate the relation between PP-based bundle frequency and writing scores, a linear regression test was applied to both high-scored and low-scored essays. Only the training set on low-scored essays yielded significant results explaining 5% of the variance with 91.27% accuracy. Since there is a strong agreement between predicted and actual scores, further studies can construct a full regression model based on the equation for low-scored essays. For example, the equation given for low-scored essays can help further studies to include more explanatory variables (e.g., lengths of lexical bundles) to better explain the relation between phraseological complexity and writing quality. However, the results of the high accuracy rate in the regression model should be approached with caution as the most essays were scored in the ranges of 5 to 7. Since the essays that are considered low-scored contain few samples of essays that range between 1 and 4, the high accuracy of the regression model might be attributable to the imbalanced dataset. Therefore, the results are only tentative as an analysis on a larger data set with better balancing of essay scores might produce different results.

There was also a decreasing trend in PP-based bundle frequency with increased writing performance scores. These findings are partially in line with the studies, which

found PP-based bundles to decrease with increased proficiency (e.g., Candarlı, 2020; Staples et al., 2013). For example, the results of this study agree with Staples et al. (2013) in that low-scored essays tend to include more tokens of PP-based lexical bundles than high-scored essays. The findings also align with the study of Candarlı (2020), which found a negative correlation between PP-based bundle frequency and increased proficiency. It is important to note that the regression model could only explain a small variance in the learner scores received between 1-6, which correspond to low-scored (1-3) and mid-scored (3-6) essays according to the rating scale used for human holistic scores. When these limitations are considered, it can tentatively be concluded that these findings partially agree with the idea that lexical bundles are essential constructs for low-proficiency learners.

The findings of the third research question partially addressed the issue mentioned by Durrant (2019) regarding the unclear link between formulaic language use and writing quality. Although the use of four-word PP-based bundles is only one aspect of overall phrasal complexity, it explained the 5% of variance in writing performance with high accuracy. There are many other potential factors signaling phraseological competence that could explain a greater proportion of variance in the scores of L2 English essays. As reported by Kyle (2016), TAASSC can quantify 132 indices of phrasal complexity. The frequency of PP-based bundles coupled with fine-grained phrasal complexity indices can help build a regression model with a better predictive power. A regression computed with the inclusion of other indices of phrasal complexity could be a venue for further research to provide conclusive evidence on the link between formulaic language use and writing quality.

5.3.2. Comparison of PP-based LBs across high- and low-scored essays

RQ3b investigated the differences in PP-based LB frequency across low-scored and high-scored essays. Contrary to previous studies (e.g., Chen & Baker, 2010, 2016; Ruan, 2017; Vo, 2019), the present study revealed no statistical difference in the use of PP-based bundle tokens and types between low-scored and high-scored essays. It is important to note that the aforementioned studies used inferential statistics such as chi-square or the log-likelihood tests. As argued by Lijffijt et al. (2016), statistical tests such as chi-square or the log-likelihood test might not be reliable when the corpus compared

is different in size and average text length. Chen and Baker (2016), for example, used a log-likelihood test, but the average text length between B1 (average length of 139 words), B2 (368 words), and C1 (559 words) was significantly different. According to Biber et al. (1998), there are more opportunities for longer texts to include more language features, thereby normalization of frequency counts is necessary for that reason. Since the Log-likelihood test is recommended to be used with raw frequencies, the reliability of the results can be affected by differences in text lengths. As argued by Pan et al. (2020), in addition to the difference in the average text length in essays, differences in the number of essays can also result in type iii errors in research. It was observed that the vast majority of the performance scores on essays in TELC were accumulated between the ranges of 5 to 7. Imbalanced distribution of the proficiency levels in the data may partly explains the non-significant differences between low-scored and high-scored essays. Additionally, both groups of students preferred other PP fragments over PP fragments with -of structure. Considering the limitation of representativeness of proficiency level, the frequency analysis across different proficiency levels tentatively confirmed that there were no significant differences between the two groups.

Aside from the studies that used chi-square or log-likelihood tests to compare bundle types and tokens, there are also studies, which compared essays across various proficiency levels using such statistical models as the Mann-Whitney U test (e.g., Kim & Kessler, 2022; Staples et al., 2013). Also, Appel (2022) broke the texts into standard size chunks before the log-likelihood analysis to address the potential confounding variables related to textual composition. Analyzing the three groups of learners at different proficiency levels based on TOEFL iBT scores, Staples et al. (2013) found low-scoring essays to include more tokens of lexical bundles. However, the bundle types differed slightly across low-scoring, mid-scoring, and high-scoring essays. While the present study contradicts the findings of Staples et al. on bundle token counts, it is important to note that this study focused solely on PP-based bundles. The higher tokens of bundles in Staples et al. might be related to the low-scoring learners' tendency to use colloquial bundles, which are generally VP-based and stance bundles in nature. Therefore, investigation of VP-based and NP-based bundles in TELC might give more comprehensive insights into the differences between high-scored and low-scored essays.

Drawing on a similar methodology, Kim and Kessler (2022) analyzed the type and token counts of 3-, 4- and 5-word lexical bundles in the argumentative essays of first and second-year university students. On a methodological note, Kim and Kessler investigated 120 essays written in a timed manner. The findings of Kim and Kessler align with the present study as they found no significant differences in the use of four-word PP-based bundles across low-scoring and high-scoring essays. However, according to their results, the most distinctive difference between these learner groups was the use of three-word bundles, as learners at higher proficiency levels used significantly more tokens and types of these bundles. A recent study by Appel (2022) found comparable results to Kim and Kessler (2022) in terms of bundle token and type distribution across different proficiency levels. Differently than Kim and Kessler (2022), Appel investigated the 3-7-word bundles in untimed essays. In Appel, high proficiency learners used more types and tokens of three-word bundles in their papers compared to learners at low proficiency. However, the type and token distribution of four-word lexical bundles did not show any statistical difference. On closer inspection, both groups used similar proportions of the bundles *on the other hand* and *one of the most*, which is in line with the present study. Appel also reported that high-scoring learner essays contained more discourse-organizing bundles, which are mostly NP-based and PP-based.

Overall, this comparative study found no significant differences in type and token frequency of 4-word PP-based bundles across low-scored and high-scored learner essays, which is in line with the recent studies (e.g., Kim & Kessler, 2022). However, the findings of the present study contradicted earlier studies that used log-likelihood or chi-square statistics but failed to control textual composition variables (e.g., text length and number of essays). Therefore, methodological differences may potentially explain conflicting results between current research and previous results of some studies (e.g., Chen & Baker, 2016; Ruan, 2017). In addition, the control of proficiency level is a limitation of the present study. A corpus with a better distribution of texts across different proficiency levels could yield more significant results. It is also worth investigating the use of bundles at different lengths and different structural categories to draw a more accurate picture of the relationship between writing proficiency and lexical bundle use.

CHAPTER 6: CONCLUSION

Lexical bundles are an essential part of academic discourse and phraseological complexity is regarded as one of the defining elements of EAP development. However, several studies claimed that L2 English learners have difficulties in the use of articles and prepositions. In addition, previous research has consistently demonstrated that NNS use fewer PP-based lexical bundles than NS writers. To shed light on this issue, Shin et al. (2018) proposed that the underuse of bundle types and tokens might be related to the nature of frequency-based corpus analysis methods. Frequency-based methods ignore the lexical bundles that contain errors, leading to the conclusion that learners use fewer tokens of lexical bundles than native speakers. One of the objectives of the present study was to test this hypothesis on PP-based lexical bundles. Contrary to Shin et al. (2018), the present study found that non-native writers use more bundles than native speakers, particularly the bundles *on the other hand* and *of the most important*. As previously argued, traditional frequency-based corpus analyses fail to account for attempted bundles as they only retrieve correct forms. Therefore, adopting a similar methodology to core expression analysis, the present study used a set of regular expression codes to extract attempted bundles. The analysis resulted in a slight increase in overall token and type counts of PP-based lexical bundles. However, the regular expression analysis revealed that *on the other hand*, which was found to be overused, is also commonly misused by Turkish learners. Appel and Murray (2020) argue that relying solely on the analysis of L1 and L2 differences does not give a comprehensive overview of the language patterns generated by L1 learners. Therefore, to further investigate use of PP-based bundles by Turkish students, the present study compared high-scored and low-scored essays using a linear regression test. There was no significant difference between the two groups of learners, indicating that overuse of certain bundles is common in both learners at lower and higher proficiency levels. However, a regression analysis with test set on low-scored essays revealed that the model can account for 5% of the variance. Also, a gradual decline in the use of PP-based bundles was observed in the essays of low-scored learners, implying that lexical bundles are essential constructs for learners at lower levels. Overall, it was found

that learners who received the lowest scores (1-4) were the ones to use lexical bundles most frequently.

6.1. Summary of findings

This thesis provided an exploratory investigation of the use of PP-based lexical bundles produced by English speaking Turkish university students, while also taking stock of frequency and accuracy of use. In a comparative analysis to investigate the first research question, there were significant differences between TELC and LOCNESS according to the Mann-Whitney U test. It was found that learners in TELC slightly overused PP-based bundles compared to LOCNESS. This finding contradicted Shin et al. (2018) and other researchers (e.g., Lee et al., 2020) who found NNS to produce fewer tokens of PP-based bundles. On a closer inspection of the occurrence of highly frequent bundles in TELC, it was found that the overuse of bundles was related to the high frequency of such bundles as *on the other hand* and *of the most important.* These two bundles together accounted for approximately half of the total bundle tokens found in TELC. Though TELC used more lexical bundle tokens, their use of bundles lacked variety compared to LOCNESS. Six bundles (*as a result, to the fact that, in the case of, on the part of, at the end of, in a way that*), most of which can be considered as framing bundles were found to be underused. The findings on underuse of bundles showed that learners in TELC have difficulties framing their opinions and using bundles that are complex in nature.

Since the quality of tokens are as important as quantity, the errors and potential uses of lexical bundles were investigated. Also, Shin et al. (2018) proposed that the lexical bundles that are erroneously used can go undetected by frequency-based methods, resulting in significantly less tokens found in NNS corpus. One of the objectives of this study was to test this hypothesis by using regular expression codes to detect attempted bundles. This study was the first to investigate the use of PP-based bundles using regular expression analysis in the Turkish context. With regards to the second question, it was found that the essays in TELC were highly accurate in terms of errors related to the use of reference bundles. Based on percentages, inaccurate bundles accounted for 5.1% of the

total number of bundles. It was found that the most common errors were related to the use of *on the other hand*, which is the most frequent PP-based bundle in TELC. These errors were related to L1 influence and misformed use of prepositions. The overuse and misuse of this bundle contributed to the arguments of L2 acquisition, which claimed that production of conventional formulaic language presents a challenge for L2 adult learners (see Yorio, 1989). As a part of the second question, regular expressions were analyzed and combined with the errors. Regular expression analysis was found to be effective as it demonstrated even in a study with limited scope (with only focusing on PP-based bundles) the analysis found 26 bundle tokens and 4 bundle types. When conducted on a larger corpus and on combination of different bundle types or lengths, the regular expression analysis has the potential to identify more attempted bundles. The attempted bundles were added to the overall list of bundles in TELC and used in the analysis of the third research question (both RQ3a and RQ3b).

As suggested by several researchers (Durrant, 2019; Kim & Kessler, 2022), the relation between the use of formulaic language and writing quality is under-researched. RQ3a examined the relationship between PP-based bundle frequency and writing quality by using a linear regression analysis. The linear regression analysis showed that total PP-based bundle frequencies were not significantly related to writing quality. Based on this, the PP-based bundle frequency does not appear to be a distinguishing feature across proficiency levels. However, it should be noted that the majority of the essays are scored between 5 to 7, with low-scored essays and high-scored essays being unevenly distributed in the corpus.

Therefore, to further explore the predictive power of PP-based bundle frequency, the dataset of each group (i.e., low- and high-scored essays) followed 80/20 principle for training and accuracy tests. The samplings for accuracy and training were created using stratified random sampling procedure to avoid overfitting the model. The initial analysis on high-scoring essays had no significant results, thereby the regression model was not tested for accuracy. Low-scored essays, on the other hand, yielded significant results and explained 5% of the variance in the overall writing quality scores with 91.27% accuracy. According to the results of the analysis, low-scored essays rated between 1 to 4 used significantly more bundles. Similarly, a longitudinal study by Candarlı (2020) revealed a

comparable pattern of PP-based bundle use, as the study found a negative correlation between PP-based bundle use and writing quality. These findings partially lend support to the idea presented in Staples et al. (2013) that formulaic language is an important construct for learners at lower proficiency levels.

RQ3b investigated the differences in the type and token distribution of PP-based bundles in low-scored and high-scored essays. To detect the patterns of overuse and underuse, a Mann-Whitney U test was performed. The test did not reveal any significant results. Although the results go against the previous studies that used log-likelihood test, those studies failed to control the confounding variables such as differences in text lengths and essay numbers. The confounding variables related to textual composition can significantly affect the results of the study and lead to type iii research errors (see Pan et al., 2020). Previous research that controlled for these confounding variables reported results similar to the findings of this study. Overall, it was found that the bundle types and tokens did not vary across the two groups of learners. The findings on the frequency of bundles across proficiency levels showed that overuse and underuse of the bundles presented in Chapter 4 was not related to the proficiency of the learners. The results regarding the RQ3b should be interpreted with caution as it is highly probable that the high proportion of essays scored between 5-6 contributed to the homogeneity of bundle counts between the two groups. Taken together, these findings lend support to three claims:

(1) Turkish learners tend to use certain PP-based bundles repetitively contributing to their overuse of bundles

(2) contrary to expectations Turkish learners were highly accurate in their use of 4-word PP-based bundles

(3) PP-based bundle frequency can predict low-scored essays.

(4) high-scoring essays or essays that can be rated mid-scoring do not necessarily contain more PP-based bundle types and tokens.

The results uncovered here can have potential pedagogical implications and contributions to the understanding of L2 writing development.

6.2. Pedagogical implications

The present study provided a unique analysis of PP-based lexical bundle use from a perspective of the accuracy of use and the relation of these bundles to writing quality. This study illustrated that Turkish learners have difficulties in the use of certain PP-based bundles. The results of this study have several implications for learners, language teaching, course and material development, and for further studies.

It was observed that Turkish learners have difficulties in the use of bundles such as *on the other hand* and *on the one hand*. As reported in the database of Collocaid (Frankenberg-Garcia & Rees, 2021), erroneous uses of these bundles are relatively common in the writings of academic English learners. An explicit formulaic language instruction (e.g., Shin & Kim, 2017) and addressing the lack of materials to teach complex syntactic phrases and clauses mentioned in Gedik and Kolsal (2022) might help learners to use these bundles more appropriately. Furthermore, the students in TELC were allowed to use reference materials such as dictionaries and other online sources. It was observed from the misuses of *on the other hand* that students preferred to use online Turkish-English dictionaries, which resulted in such uses as *on the other side* and *on the other way* as an equivalent to the intended bundle. Although there were few uses such as those, it points to the interference of L1. To mitigate the potential influence of Turkish on language learning, the use of concordancers can be encouraged to promote learner autonomy and more appropriate use of bundles.

Considering that some PP-based lexical bundles are used excessively, and Turkish students exhibited lack of variety in their use of bundles, it is crucial to raise students' awareness of their frequency, functions, and appropriate use. Focused explicit instruction on formulaic language can promote the acquisition and internalization of these structures, leading to better writing proficiency (see Alhassan, 2018). For example, Alhassan used controlled rewriting teaching activities and free writing activities with salient input to increase learners' awareness of formulaic language. The worksheets used in Alhassan contain activities related to the types of bundles that were overused, underused, and misused in this study. They can serve as an effective material for teaching both alternative

structures for overused bundles and appropriate uses of PP-based bundles that were found to be commonly misused.

The findings also have tentative implications for course development. The present study has also provided clear evidence that PP-based bundle frequency is a significant predictor of writing quality in low-scored essays. In Türkiye, there is a mismatch between the official language education materials and nationwide exams in terms of phraseological complexity (Gedik & Kolsal, 2022). The official English coursebooks in Türkiye are not rich enough in syntactic and lexical complexity to prepare students for tertiary education. Including more PP-based bundles in English coursebooks can help students better prepare for high-stakes exams since PP-based bundles have been found to predict writing quality and are typical of academic writing. In addition, the official coursebooks can include clues or tip boxes that inform students about the erroneous uses of bundles (e.g., *in the other hand*). Such prompts can help students to organize their arguments more effectively and increase their awareness of the alternative constructs for the overused bundles.

The findings on the relationship between writing quality and PP-based lexical bundle frequency provided insights into usage-based theories. It was shown that PP-based lexical bundles are a challenge to Turkish adult L2 learners, as they were found to overuse certain bundles. In addition, the writing quality negatively correlated with PP-based lexical bundle frequency in low-scored essays. It indicated that though bundles were employed in low-scored essays, they might have been misused semantically or overused as they negatively affected the scores given by human raters. Moreover, the regression equation given for low-scored essays can help further studies construct a full regression model. Last but not least, language research can benefit from the regular expression code sheet prepared for PP-based reference bundles.

6.3. Further research and limitations

There are several limitations to acknowledge in this study. Overall, the regular expression method produced meaningful results, but there is a need for further studies on different types of bundles. In this study, only PP-based bundles were analyzed, so

considering more bundle types might draw a more accurate picture of Turkish EFL learners' use of lexical bundles. Therefore, in further studies, it would be of great interest to investigate the use of NP- or VP-based four-word reference bundles to test the effectiveness of using regular expression codes in error analysis. Also, since the findings revealed that many bundles appeared only once or twice, investigating 4-word PP-based bundles in larger datasets holds the potential to produce more meaningful results in terms of frequency counts. It is possible to identify more errors in a larger corpus, and with a larger sample of errors, it would have been possible to conduct a linear regression analysis to investigate the relation of errors to writing performance. For the relation between writing performance and PP-based bundle frequency, a more balanced dataset could yield more reliable results.

REFERENCES

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. https://doi.org/10.1016/j.esp.2011.08.004

Anthony, L. (2022). AntConc (Version 4.2.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

AlHassan, L. (2018). *An Empirical Investigation of the Role of Formulaic Sequences in Upgrading EAP Students' Academic Writing Skills* [Doctoral dissertation, Carleton University]. The Carleton University Institutional Repository. https://repository.library.carleton.ca/concern/etds/br86b4576

Allen, D. (2009). Lexical bundles in learner writing: An analysis of formulaic language in the ALESS learner corpus. *Komaba Journal of English Education*, *1*(4), 105-127.

Altenberg, B. (1991). 'A bibliography of publications relating to English computer corpora'. In S. Johansson & A. B. Stenstrom (eds.), *English computer corpora: Selected papers and research guide* (pp. 355-396). Mouton https://doi.org/10.1515/9783110865967.355

Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 101–122). Oxford University Press.

Altınmakas, D., & Bayyurt, Y. (2019). An exploratory study on factors influencing undergraduate students' academic writing practices in Turkey. *Journal of English for Academic Purposes, 37*, 88-103. https://doi.org/10.1016/j.jeap.2018.11.006

Appel, R. (2022). Lexical bundles in L2 English academic texts: relationships with holistic assessments of writing quality. *System*, *110*, 102899. https://doi.org/10.1016/j.system.2022.102899

Appel, R., & Murray, L. (2020). L1 differences in L2 English academic writing: A lexical bundles analysis. *Journal of English for Academic Purposes*, *46*, 100873. https://doi.org/10.1016/j.jeap.2020.100873

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. https://doi.org/10.1080/15434303.2015.1126718

Bal, B. (2010). *Analysis of four-word lexical bundles in published research articles written by Turkish scholars* [Master's thesis, Georgia State University]. Retrieved from http://scholarworks.gsu.edu/alesl_theses/2

Becker, J. D. (1975). The phrasal lexicon. In Y. Wilks (Ed.), TINLAP '75. *Proceedings of the 1975 workshop on Theoretical issues in natural language processing* (pp. 60–63). Association for Computational Linguistics https://doi.org/10.3115/980190.980212

Bestgen, Y. (2017). Beyond single-word measures: L2 writing assessment, lexical richness, and formulaic competence. *System*, *69*, 65–78. https://doi.org/10.1016/j.system.2017.08.004

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. https://doi.org/10.1016/j.jslw.2014.09.004

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers.* John Benjamins. https://doi.org/10.1075/scl.23

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for specific purposes*, *26*(3), 263-286. https://doi.org/10.1016/j.esp.2006.08.003

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at…: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*(3), 371-405. https://doi.org/10.1093/applin/25.3.371

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press. https://doi.org/10.1017/CBO9780511804489

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development?. *Tesol Quarterly*, *45*(1), 5-35. https://doi.org/10.5054/tq.2011.244483

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

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, *84*(4), 711-733. https://doi.org/10.1353/lan.2006.0186

Bychkovska, T., & Lee, J. J. (2017). At the same time: Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for Academic Purposes*, *30*, 38-52. https://doi.org/10.1016/j.jeap.2017.10.008

Campion, M., & Elley, W. (1971). *An academic vocabulary list*. New Zealand Council for Educational Research.

Cangır, H. (2018). *Investigating the Relationship between L1 and L2 Collocational Processing in the Bilingual Mental Lexicon* [Doctoral dissertation, Hacettepe University]. http://www.openaccess.hacettepe.edu.tr:8080/xmlui/handle/11655/4986

Chen, A. C. H. (2019). Assessing phraseological development in word sequences of variable lengths in second language texts using directional association measures. *Language learning*, *69*(2), 440-477. https://doi.org/10.1111/lang.12340

Chen, Y. (2009) *Investigating Lexical Bundles Across Learner Writing Development*. [Doctoral dissertation, Lancaster University]. EThOS.

Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing.

Language Learning and Technology, *14*(2), 30–49. https://doi.org/10.10125/44213

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.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on information theory*, *2*(3), 113-124. https://doi.org/10.1109/TIT.1956.1056813

Chomsky, N. (1967). Preface to 'A Review of B. F. Skinner's Verbal Behavior'. In L. A. Jakobovits and M. S. Miron (eds.), *Readings in the Psychology of Language* (pp. 142–143). Prentice-Hall.

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. https://doi.org/10.1093/applin/amm022

Cooper, T. (2013). Can IELTS writing scores predict university performance? Comparing the use of lexical bundles in IELTS writing tests and first-year academic writing. *Stellenbosch Papers in Linguistics Plus*, *42*, 63-79. http://dx.doi.org/10.5842/42-0-155

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, *23*(4), 397-423. https://doi.org/10.1016/j.esp.2003.12.001

Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes*, *12*, 33–43. https://doi.org/10.1016/j. jeap.2012.11.002

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*, 213-238. https://doi.org/10.2307/3587951

Crossley, S. A. (2020). Linguistic features in writing quality and development: An

overview. *Journal of Writing Research*, *11*(3), 415–443. https://doi.org/10.17239/jowr-2020.11.03.01

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing*, *32*, 1-16. https://doi.org/10.1016/j.jslw.2016.01.003

Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading*, *35*(2), 115–135. https://doi.org/10.1111/j.1467-9817.2010.01449.x

Csomay, E. (2013). Lexical bundles in discourse structure: A corpus-based study of classroom discourse. *Applied linguistics*, *34*(3), 369-388. https://doi.org/10.1093/applin/ams045

Dalkey, N., & Helmer, O. (1963). An experimental application of the DELPHI method to the use of experts. *Management Science*, *9*(3), 458–467. https://doi.org/10.1287/mnsc.9.3.458

Demirel, T. E. (2017). Detection of common errors in Turkish EFL students' writing through a corpus analytic approach. *English Language Teaching*, *10*(10), 159-178. https://doi.org/10.5539/elt.v10n10p159

Dulay, H., Burt, M., & Krashen, S. D. (1982). *Language two*. Oxford University Press

Durrant, P. (2019). Formulaic language in English for academic purposes. In A. Siyanova-Chanturia, & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (1st ed., pp. 211-227). Routledge. https://doi.org/10.4324/9781315206615-12

Durrant, P., & Mathews-Aydinli, J. (2011). A function-first approach to identifying formulaic language in academic writing. *English for Specific Purposes*, *30*(1), 58-72. https://doi.org/10.1016/j.esp.2010.05.002

Ellis, N. C., Simpson-Vlach, R. I. T. A., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, *42*(3), 375-396. https://doi.org/10.1002/j.1545-7249.2008.tb00137.x

Ellis, R. (2014). Principles of instructed second language learning. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (4th ed., pp. 31-45). National Geographic Learning. https://doi.org/10.1016/j.system.2004.12.006

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, *20*(1), 29–62. https://doi.org/10.1515/text.1.2000.20.1.29

Ertürk, G. T., & Öztürk, K. (2022). Mixed-Method Research on EFL Graduate Students' Academic Writing Practices. *English Language Teaching*, *15*(7), 110-126. https://doi.org/10.5539/elt.v15n7p110

Esfandiari, R., Barbary, F. (2017). A contrastive corpus-driven study of lexical bundles between English writers and Persian writers in psychology research articles. *Journal of English for Academic Purpose*, *29*, 21-42. https://doi.org/10.1016/j.jeap.2017.09.002

Firth, J. R. (1957). A synopsis of linguistic theory, 1930–55. In *Studies in Linguistic Analysis* (pp. 1–31). Special Volume of the Philological Society. Blackwell [Reprinted as Firth (1968)].

Flowerdew, J. (2019). The linguistic disadvantage of scholars who write in English as an additional language: Myth or reality. *Language Teaching*, *52*(2), 249-260. https://doi.org/10.1017/S0261444819000041

Frankenberg-Garcia, A., & Rees, G. (2021). ColloCaid Academic Collocation Errors and Other Problems (Version 1). figshare. https://doi.org/10.6084/m9.figshare.13640624.v1

Garner, J., Crossley, S. A., & Kyle, K. (2019). N-gram measures and L2 writing

proficiency. *System*, *80*, 176-187. https://doi.org/10.1016/j.system.2018.12.001

Gass, S., Behney, J., & Plonsky, L. (2014). *Second language acquisition: An introductory course.* Routledge. https://doi.org/10.4324/9780203137093

Gedik, T. A., & Kolsal, Y. S. (2022). A corpus-based analysis of high school English textbooks and English university entrance exams in Turkey. *Theory and Practice of Second Language Acquisition*, *8*(1), 157-176. https://doi.org/10.31261/TAPSLA.9152

Geluso, J. (2019). *Frequency, semantic, and functional characteristics of discontinuous formulaic language: A learner corpus study* (Publication No. 17682). [Doctoral dissertation, Iowa State University]. Iowa State University Digital Repository.

Geluso, J. (2022). Grammatical and functional characteristics of preposition-based phrase frames in English argumentative essays by L1 English and Spanish speakers. *Journal of English for Academic Purposes*, *55*, 101072. https://doi.org/10.1016/j.jeap.2021.101072

Granger, S. (1993). International Corpus of Learner English. In Aarts, J., de Haan, P., & Oostdijk, N. (eds.) *English language corpora: Design, analysis and exploitation*, (pp. 57 – 71). Rodopi.

Granger, S. (1996). From CA to CIA and back: An integrated contrastive approach to computerized bilingual and learner corpora. In K. Aijmer, B. Altenberg, & M. Johansson (eds.), *Languages in Contrast: Papers from a symposium of text-based cross-linguistic studies* (pp. 37–51). Lund University Press.

Granger, S. (1998). The computer learner corpus: A versatile new source of data for SLA research. In S. Granger (ed.), *Learner English on computer* (pp. 3-18). Addison Wesley Longman. https://doi.org/10.4324/9781315841342-1

Granger, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching. *Corpora and language teaching*, *33*, 13-32. https://doi.org/10.1075/scl.33.04gra

Granger, S. (2015a). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research*, *1*(1), 7-24. https://doi.org/10.1075/ijlcr.1.1.01gra

Granger, S. (2015b). The contribution of learner corpora to reference and instructional materials design. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 485-510). Cambridge University Press. https://doi.org/10.1017/cbo9781139649414.022

Granger, S., Dupont, M., Meunier, F., Naets, H., & Paquot, M. (2020). *The International Corpus of Learner English. Version 3*. Presses universitaires de Louvain.

Gilquin, G. (2022). Cognitive corpus linguistics and pedagogy: From rationale to applications. *Pedagogical Linguistics*, 3(2), 109-142. https://doi.org/10.1075/pl.22014.gil

Güngör, F., & Uysal, H. H. (2016). A comparative analysis of lexical bundles used by native and non-native scholars. *English Language Teaching, 9*(6), 176-188. https://doi.org/10.5539/elt.v9n6p176

Hasselgård, H. (2019). Phraseological teddy bears: Frequent lexical bundles in academic writing by Norwegian learners and native speakers of English. In M. Mahlberg, & V. Wiegand (Eds.), *Corpus Linguistics, context and culture* (pp. 339-362). De Gruyter. https://doi.org/10.1515/9783110489071-013

Hasselgård, H., & Johansson, S. (2011). Learner corpora and contrastive interlanguage analysis. In F. Meunier, S. De Cock, G. Gilquin, & M. Paquot (Eds.), *A Taste for Corpora: In honour of Sylviane Granger* (pp. 33–62). John Benjamins. https://doi.org/10.1075/scl.45.06has

Hejazi, H. (2021). *A Corpus-Based Investigation of Lexical Bundles and Keyness in B1, B2 and C1 ESL Learners' Academic Writing* [Doctoral dissertation, University of Liverpool]. University of Liverpool. The University of Liverpool Repository. https://livrepository.liverpool.ac.uk/id/eprint/3154170

Huang, K. (2015). More does not mean better: Frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing. *System, 53*, 13–23. https://doi.org/10.1016/j.system.2015.06.011

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes, 27*(1), 4-21. https://doi.org/10.1016/j.esp.2007.06.001

Hyland, K. (2012). Bundles in academic discourse. *Annual review of applied linguistics, 32*, 150-169. https://doi.org/10.1017/S0267190512000037

Hyland, K., & Tse, P. (2007). Is there an "academic vocabulary"? *TESOL Quarterly, 41*(2), 235-253. https://doi.org/10.1002/j.1545-7249.2007.tb00058.x

Ishikawa, S. (2011). Phraseology overused and underused by Japanese learners of English. In K. Yagi, T. Kanzaki, & A. Inoue (Eds.), *Phraseology, corpus linguistics and lexicography: Papers from Phraseology 2009 in Japan* (pp. 82-94). Kwansei Gakuin University Press.

Kashiha, H., & Chan, S. H. (2015) A little bit about: Differences in native and non-native speakers' use of formulaic language. *Australian Journal of Linguistics, 35*(4), 297–310. https://doi.org/10.1080/0726 8602.2015.1067132

Kellogg, R. T., & Raulerson, B. A. (2007). Improving the writing skills of college students. *Psychonomic Bulletin & Review*, *14*(2), 237–242. https://doi.org/10.3758/BF03194058

Kim, S., and Kessler, M. (2022). Examining L2 English university students' uses of lexical bundles and their relationship to writing quality. *Assessing Writing, 51*, 100589. https://doi.org/10.1016/j.asw.2021.100589

Köse, G. D., Yüksel, İ., Öztürk, Y., & Tömen, M. (2019). Turkish Academics' Foreign Language Academic Literacy: A Needs Analysis Study. *International Journal of Instruction, 12*(1), 717-736. https://doi.org/10.29333/iji.2019.12146a

Kyle, K. (2016). Measuring syntactic development in L2 writing: Fine grained indices of

syntactic complexity and usage-based indices of syntactic sophistication [Doctoral dissertation, Georgia State University]. ScholarWorks @Georgia State University. http://scholarworks.gsu.edu/alesl_diss/35

Kyle, K., & Crossley, S. A. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *Modern Language Journal*, *102*(2), 333–349. https://doi.org/10.1111/modl.12468

Langacker, R. W. (1987). *Foundations of cognitive grammar: Volume I: Theoretical prerequisites* (Vol. 1). Stanford university press.

Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow, & S. Kemmer (Eds.), *Usage-based model of language* (pp. 24-63). Stanford University Press.

Lee, Y. E., Yoo, I. W. H., & Shin, Y. K. (2020). The use of English prepositions in lexical bundles in essays written by Korean university students. *Journal of English for Academic Purposes*, *45*, 100848. https://doi.org/10.1016/j.jeap.2020.100848

Lu, X., & Deng, J. (2019). With the rapid development: A contrastive analysis of lexical bundles in dissertation abstracts by Chinese and L1 English doctoral students. *Journal of English for Academic Purposes*, *39*, 21-36. https://doi.org/10.1016/j.jeap.2019.03.008

McEnery, T., Xiao, R. and Tono, Y. (2006). *Corpus-Based Language Studies. An Advanced Resource Book*. Routledge.

McNamara, D. S., Crossley, S. A., & McCarthy, P.M. (2010). Linguistic features of writing quality. *Written Communication*, *27*(1), 57–86. https://doi.org/10.1177/0741088309351547

Musgrave, J., & Parkinson, J. (2014). Getting to grips with noun groups. *ELT Journal*, *68*(2), 145-154. https://doi.org/10.1093/elt/cct078

Nation, P. (1990). *Teaching and learning vocabulary*. Newbury House

Pinto, P. T., Crosthwaite, P., de Carvalho, C. T., Spinelli, F., Serpa, T., Garcia, W., &

Ottaiano, A. O. (Eds.). (2023). *Using Language Data to Learn About Language: A Teachers' Guideto Classroom Corpus Use.* University of Queensland. https://doi.org/10.14264/3bbe92d

Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for Academic Purposes*, *21*, 60-71. https://doi.org/10.1016/j.jeap.2015.11.003

Pan, F., Reppen, R., & Biber, D. (2020). Methodological issues in contrastive lexical bundle research: The influence of corpus design on bundle identification. *International Journal of Corpus Linguistics*, *25*(2), 216-230. https://doi.org/10.1075/ijcl.19063.pan

Paquot, M. (2010). *Academic vocabulary in learner writing: From extraction to analysis*. Continuum.

Paquot, M. (2018). Phraseological competence: A missing component in university entrance language tests? Insights from a study of EFL learners' use of statistical collocations. *Language Assessment Quarterly*, *15*(1), 29-43. https://doi.org/10.1080/15434303.2017.1405421

Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, *32*, 130-149. https://doi.org/10.1017/S0267190512000098

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. 191–227). Longman.

Praninskas, J. (1972). *American university word list*. Longman.

Qin, J. (2014). Use of formulaic bundles by non-native English graduate writers and published authors in applied linguistics. *System*, *42*, 220-231. https://doi.org/10.1016/j.system.2013.12.003

Ruan, Z. (2017). Lexical bundles in Chinese undergraduate academic writing at an English medium university. *RELC Journal*, *48*(3), 327-340. https://doi.org/10.1177/0033688216631218

Schmitt, N. (2005). Formulaic language: Fixed and varied. *Estudios de Lingüística Inglesa Aplicada*, *6*, 13–39.

Schmitt, N. (2022). Norbert Schmitt's essential bookshelf: Formulaic language. *Language Teaching*, 1-12. https://doi.org/10.1017/S0261444822000039

Schmitt, N., & Carter, R. (2004). Formulaic Sequences in Action. In N. Schmitt (ed.), *Formulaic Sequences Acquisition, processing and use (pp. 1-22).* John Benjamins Publishing Company. https://doi.org/10.1075/lllt.9

Shin, Y. K. (2018). *Lexical bundles in argumentative essays by native and nonnative English-speaking novice academic writers* [Doctoral dissertation, Georgia State University]. ScholarWorks @Georgia State University. https://scholarworks.gsu.edu/alesl_diss/47/

Shin, Y. K. (2019). Do native writers always have a head start over nonnative writers? The use of lexical bundles in college students' essays. *Journal of English for Academic Purposes*, *40*, 1-14. https://doi.org/10.1016/j.jeap.2019.04.004

Shin, Y. K., & Kim, Y. (2017). Using lexical bundles to teach articles to L2 English learners of different proficiencies. *System*, *69*, 79-91. https://doi.org/10.1016/j.system.2017.08.002

Shin, Y. K., Cortes, V., & Yoo, I. W. (2018). Using lexical bundles as a tool to analyze definite article use in L2 academic writing. *Journal of Second Language Writing*, 39, 29–41. https://doi.org/10.1016/j.jslw.2017.09.004

Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied linguistics*, *31*(4), 487-512. https://doi.org/10.1093/applin/amp058

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, *27*(2), 251-272. https://doi.org/10.1177/0267658310382068

Sosa, A. V., & MacFarlane, J. (2002). Evidence for frequency-based constituents in the mental lexicon: Collocations involving the word of. *Brain and Language*, *83*(2), 227–236. https://doi.org/10.1016/S0093-934X(02)00032-9

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*(3), 214-225. https://doi.org/10.1016/j.jeap.2013.05.002

Stubbs, M. (2004). Language corpora. In A. Davies, & C. Elder (Eds.), *The Handbook of applied linguistics* (pp. 106-132). Blackwell. https://doi.org/10.1002/9780470757000.ch4

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics, 11*, 61 – 82. https://doi.org/10.1515/cogl.2001.012

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

Torlak, M. (2020). *Investigating the Role of Multi-Word Expressions in Writing Proficiency and Overall English Language Proficiency of EFL Students at a Turkish University*. [Master's thesis, Bilkent Üniversitesi]. ProQuest Dissertations and Theses Global.

Ucar, S. (2017). A Corpus-Based Study on the Use of Three-Word Lexical Bundles in the Academic Writing by Native English and Turkish Non-Native Writers. *English Language Teaching*, *10*(12), 28-36. https://doi.org/10.5539/elt.v10n12p28

Underhill, R. (1976). *Turkish grammar*. MIT Press.

Uzun, K. (2018). The use of lexical bundles and the definite article 'the': A core expression analysis. *Cumhuriyet International Journal of Education*, *7*, 269–286. https://doi.org/10.30703/cije.441596

Vo, S. (2019). Use of lexical features in non-native academic writing. *Journal of Second Language Writing*, *44*, 1-12. https://doi.org/10.1016/j.jslw.2018.11.002

Weinert, R. (1995). The role of formulaic language in second language acquisition: A review. *Applied linguistics*, *16*(2), 180-205. https://doi.org/10.1093/applin/16.2.180

Wood, D. (2015). *Fundamentals of formulaic language: An introduction*. Bloomsbury Publishing.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press. https://doi.org/10.1017/CBO9780511519772

Yoon, C., & Choi, J. M. (2015). Lexical bundles in Korean university students' EFL compositions: A comparative study of register and use. *Modern English Education*, *16*(3), 47-69. https://doi.org/10.18095/meeso.2015.16.3.03

Yoon, H. J. (2021). L2 Writing and Formulaic Language: Formulaic Chunks and Lexical Bundles. In *The Routledge Handbook of Second Language Acquisition and Writing* (pp. 199-212). Routledge. https://doi.org/10.4324/9780429199691-22

Yorio, C. A. (1989). Idiomaticity as an indicator of second language proficiency. In K. Hyltenstam, & L. K. Obler (Eds.), *Bilingualism Across the Lifespan: Aspects of Acquisition, Maturity and Loss* (pp. 55–72). Cambridge University Press. https://doi.org/10.1017/CBO9780511611780.005

APPENDICES

Appendix 1: Regular Expression Code Set for PP-based Bundles

Number	Lexical Bundles	Regular Expressions
1	(About/of) the nature of	[^about|^of] (\w|\W) ^nature of$ (^about$\|^of$) (\w|\W) ^nature$ [^of]
2	as a result of	[^as] (\w|\W) (^result$\|^results$) ^of$ ^as (\w|\W) (^result$\|^results$) [^of]
3	as a function of	[^as] (\w|\W) ^function of$ ^as (\w|\W) function$ [^of]
4	as part of (a/the)	[^as] (^part$\|^parts$) ^of$ (\w|\W) ^as$ (^part$\|^parts$) [^of] (\w|\W)
5	on the part of	[^on] (\w|\W) (^part$\|^parts$) ^of$ ^on$ (\w|\W) (^part$\|^parts$) [^of]
6	as a consequence of	[^as] (\w|\W) ^consequence$ ^of$ ^as$ (\w|\W) ^consequence$ [^of]
7	as a matter of	[^as] (\w|\W) ^matter of$ ^as$ (\w|\W) ^matter$ [^of]
8	as a means of	[^as] (\w|\W) (^means$\|^mean$) ^of$ ^as$ (\w|\W) (^means$\|^mean$) [^of]
9	(at/by) the end of	[^at|^by] (\w|\W) ^end$ ^of$ (^at$\|by$) (\w|\W) ^end$ [^of]
10	at the time of	[^at] (\w|\W) (^time$\|^times$) ^of$ ^at$ (\w|\W) (^time$\|^times$) [^of]
11	at the beginning of	[^at] (\w|\W) ^beginning$ ^of$ ^at$ (\w|\W) ^beginning$ [^of]
12	at the level of	[^at] (\w|\W) (^level$\|^levels$) ^of$ ^at$ (\w|\W) (^level$\|^levels$) [^of]
13	at the expense of	[^at] (\w|\W) ^expense$ ^of$ ^at$ (\w|\W) ^expense$ [^of]
14	to the start of	[^to] (\w|\W) ^start$ ^of$ ^to$ (\w|\W) ^start$ [^of]
15	(by/in/to) the presence of	[^by$\|^in$|^to$] (\w\|\W) ^presence$ ^of$ (^by$\|^in$|^to$) (\w\|\W) ^presence$ [^of]
16	(for/in/to) the development of	[^for$\|^in$|^to$] (\w\|\W) ^development$ ^of$ (^for$\|^in$|^to$) (\w\|\W) ^development$ [^of]
17	for the purpose(s) of	[^for] (\w|\W) (^purpose$\|^purposes$) of$ ^for$ (\w|\W) (^purpose$\|^purposes$) [^of]
18	from the point of	[^from] (\w|\W) ^point$ ^of$ ^from$ (\w|\W) ^point$ [^of]

Appendix 1. Continued

Number	Lexical Bundles	Regular Expressions
19	in the case of	[^in] (\w|\W) (^case$\|^cases$) ^of$ ^in$ [^that|^this] (^case$\|^cases$) [^of]
20	as in the case	^as$ [^in] (\w|\W) (^case$\|^cases$)
21	in this case the	[^in] (^that$\|^this$|\W) (^case$\|^cases$) (\w|\W) ^in$ [^this|^the] (^case$\|^cases$) (\w|\W) ^in$ (^that$\|^this$|\W) (^case$\|^cases$) [^the]
22	in the absence of	[^in] (\w|\W) ^absence of$ ^in$ (\w|\W) ^absence [^of]
23	in the form of	[^in] (\w|\W) (^form$\|^forms$) ^of$ ^in$ (\w|\W) (^form$\|^forms$) [^of]
24	(in / a) a number of, in the number of	[^of|^in] (\w|\W) ^number$ ^of$ (^in$\|^of$) (\w|\W) ^number$ [^of]
25	in terms of the	[^in] ^terms$ ^of$ (\w|\W) ^in$ ^terms$ [^of] (\w|\W) ^in$ ^terms$ ^of$ [^the]
26	in the context of	[^in] (\w|\W) ^context ^of$ ^in$ (\w|\W) ^context$ [^of]
27	in the course of	[^in] (\w|\W) ^course$ ^of$ ^in$ (\w|\W) ^course$ [^of]
28	in the process of	[^in] (\w|\W) ^process$ ^of$ ^in$ (\w|\W) ^process$ [^of]
29	in a variety of	[^in] (\w|\W) ^variety$ ^of$ ^in$ (\w|\W) ^variety$ [^of]
30	in the area of	[^in] (\w|\W) (^area$\|^areas$) ^of$ ^in$ (\w|\W) (^area$\|^areas$) [^of]
31	in the direction of	[^in] (\w|\W) ^direction$ ^of$ ^in$ (\w|\W) ^direction$ [^of]
32	in the face of	[^in] (\w|\W) ^face$ ^of$ ^in$ (\w|\W) ^face$ [^of]
33	in the formation of	[^in] (\w|\W) ^formation of$ ^in$ (\w|\W) ^formation [^of]
34	in the pathogenesis of	[^in] (\w|\W) ^pathogenesis of$ ^in$ (\w|\W) ^pathogenesis [^of]
35	in the study of	[^in] (\w|\W) ^study of$ ^in$ (\w|\W) ^study [^of]
36	in the treatment of	[^in] (\w|\W) ^treatment$ ^of$ ^in$ (\w|\W) ^treatment$ [^of]
37	(in / of / to / by) the use of	[^in|^of|^to|^by] (\w|\W) ^use$ ^of$ (^in$\|^of$|^to$\|^by$) (\w|\W) ^use$ [^of]
38	in view of the	[^in] ^view$ ^of$ (\w|\W) ^in$ ^view$ [^of] (\w|\W)

Appendix 1. Continued

Number	Lexical Bundles	Regular Expressions
39	of some of the	[^of] ^some$ ^of$ (\w|\W) ^of$ ^some$ [^of] (\w|\W)
40	of the effects of	[^of] (\w|\W) (^effect$\|^effects$) ^of$ ^of$ (\w|\W) (^effect$\|^effects$) [^of]
41	on the basis of	[^on] (\w|\W) ^basis$ ^of$ ^on$ (\w|\W) ^basis$ [^of]
42	on the surface of	[^on] (\w|\W) (^surface$\|^surfaces$) ^of$ on$ (\w|\W) (^surface$\|^surfaces$) [^of]
43	over a period of	[^over] (\w|\W) ^period$ ^of$ ^over$ (\w|\W) ^period$ [^of]
44	to that of the	[^to] ^that$ ^of$ (\w|\W) ^to$ ^that$ [^of] (\w|\W)
45	at the same time	[^at] (\w|\W) same time ^in$ addition$ [^to] (\w|\W)
46	between the two groups	[^between] (\w|\W) ^two$ ^groups$
47	(by / to / from / of) the fact that	[^by|^to|^from|^of] (\w|\W) (^fact$\|^facts$) ^that$
48	for the first time	[^for] (\w|\W) ^first$ ^time$
49	in such a way	[^in] ^such$ (\w|\W) ^way$
50	in the same way	[^in] (\w|\W) ^same$ ^way$
51	in the present study	[^in] (\w|\W) ^present$ ^study$
52	in a way that	[^in] (\w|\W) ^way$ ^that$
53	in addition to the	[^in] ^addition$ ^to$ (\w|\W) ^in$ addition$ [^to] (\w|\W)
54	in an attempt to	[^in] (\w|\W) ^attempt$ ^to$ ^in$ (\w|\W) ^attempt$ [^to]
55	in contrast to the	[^in] ^contrast$ ^to$ (\w|\W) ^in$ ^contrast$ [^to] (\w|\W)
56	in relation to the	[^in] ^relation$ ^to$ (\w|\W) ^in$ ^relation$ [^to] (\w|\W)
57	in the early stages	[^in] (\w|\W) ^early$ (^stage$\|^stages$)
58	in the first place	[^in] (\w|\W) ^first$ ^place$
59	in the next (chapter / section)	[^in] (\w|\W) next (^section$\|^chapter$)
60	in the sense that	[^in] (\w|\W) ^sense$ ^that$
61	of the most important	[^of] (\w|\W) ^most$ ^important$
62	on the (other / one) hand	[^on] (\w|\W) ^(^other$\|^one$) ^hand$ ^on$ (\w|\W) (^one$\|^other$) [^hand]
63	on the grounds that	[^on] (\w|\W) ^grounds$ ^that$
64	similar to (that / those) of	^similar$ [^to$] (^that$|^those$) ^of$ ^similar$ ^to$ (^that$\|^those$) [^of]
65	to the extent that	[^to] (\w|\W) ^extent$ ^that$
66	with respect to the	[^with] ^respect$ ^to$ (\w|\W) ^with$ respect$ [^to] (\w|\W)
	Total: 86 Bundles	**Total:**119 Regex codes

Appendix 2: Essay Scoring Rubric

	Weak (0.5)	Needs Improvement (1.0)	Good (1.5)	Strong (2.0)
VOCABULARY	Predominantly high-frequency, contextually inappropriate, non-natural vocabulary and collocations. High frequency of repetitions without using synonyms. Frequent lexical and/or derivation errors.	Frequent high-frequency, contextually inappropriate, non-natural vocabulary and collocations with some deviations from the topic. Certain parts with repetitions. More than a few lexical or derivation errors.	Mainly low-frequency, contextually appropriate, natural vocabulary and collocations related to the topic with a few deviations. Generally avoids repetition with synonyms. A few lexical or derivation errors.	Predominantly low-frequency, contextually appropriate, natural vocabulary and collocations related to the topic. Avoids repetition with good use of synonyms. Little or no lexical or derivation errors.
GRAMMAR	A lack of appropriate, semantically-accurate, complex structures and affixation used in academic writing. Erroneous, distractingly repetitive, too short or too long sentences making the text difficult to understand and causing communication breakdown. Frequent tense and aspect errors.	Frequent inappropriate, semantically-inaccurate, simple structures and affixation that deviate from academic writing conventions. Frequent sentences that are too long or too short making the text difficult to read at times. Some communication breakdowns. More than a few tense and aspect errors.	Generally appropriate, semantically-accurate, complex structures and affixation used in academic writing with a few deviations. Generally appropriate sentence length. A few tense and aspect errors.	A wide range of appropriate, semantically-accurate, complex structures and affixation used in academic writing. Appropriate sentence length. Little or no tense and aspect error.
TASK COMPLETION	Unclear introduction, main body and conclusion with no thesis statement. Does not present, support and conclude arguments in the main body or develops them weakly in a run-on manner in the same paragraph. Does not consolidate the thesis in the conclusion. Irrelevant to the topic, unable to reach the word limits or neatness requires attention.	Visible clarity problems with introduction, main body and conclusion. Weak thesis statement. Clarity problems with presentation, support and conclusion for arguments that may be run-on in the same paragraph. Consolidation is present but not well-connected to the rest of the essay. Content irrelevant to the topic at times, a bit far off of the word limits and is not neat.	Generally clear introduction, main body and conclusion with a thesis statement to develop arguments from. Main body arguments are mostly related to the thesis and a paragraph is dedicated per argument with slight deviations. The thesis is consolidated in the conclusion. Mostly relevant to the topic, around the word limits, sufficiently neat and assembled with a professional look.	A clear introduction, main body and conclusion. Includes a thesis statement with a few points to develop arguments from. Presents, supports and concludes arguments in the main body directly related to the thesis. A paragraph is dedicated per argument. The thesis is consolidated in the conclusion. Relevant to the topic, within the word limits, neat and correctly assembled with a professional look.
COHERENCE – COHESION	Conceptually unrelated paragraphs that do not form a whole. Ideas do not flow in a logical order, marked by serious disorganisation in their connection. Thought pattern exists but difficult to follow. A lack of appropriate connectives, pronouns and articles.	Paragraphs only slightly related and loosely form a whole. Many ideas do not flow in a logical order, causing disorganisation in many parts. Thought pattern loosely expressed. Frequent problems in the use of connectives, pronouns and articles.	Generally related paragraphs that form a whole. Ideas mostly clear, flow in a logical order and are connected coherently with only a few deviations. Connectives, pronouns and articles are mostly used appropriately with only a few problems.	Conceptually related paragraphs that form a whole. Ideas clearly presented, flow in a logical order and are connected coherently. Appropriate use of connectives, pronouns and articles where necessary.
SPELLING – PUNCTUATION	Frequent spelling, punctuation or capitalization errors which distract the reader and make the text difficult to read.	More than a few spelling, punctuation or capitalization errors. More than a few problems with the use of conventions that may partially reduce readability.	A few spelling, punctuation or capitalization errors. A few problems with the use of conventions that may slightly reduce readability.	Little or no spelling, punctuation or capitalization error. Conventions effectively used to enhance readability.

Appendix 3: Approval of Research Committee

Evrak Tarih Sayısı: 17.03.2023-423536

T.C.
TRAKYA ÜNİVERSİTESİ REKTÖRLÜĞÜ
Sosyal ve Beşeri Bilimler Araştırmaları Etik Kurulu Başkanlığı

Sayı : E-29563864-050.03.04-423536 17.03.2023
Konu : Kararlar

Sayın Doç. Dr. Kutay UZUN

Danışmanlığını yaptığınız, Trakya Üniversitesi Sosyal Bilimler Enstitüsü Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili ve Eğitimi Yüksek Lisans Programı öğrencisi Ömer Faruk KAYA tarafından, Trakya Üniversitesi Sosyal ve Beşeri Bilimler Araştırmaları Etik Kurulu'nun 15.03.2023 tarihli toplantısında alınan 02/05 numaralı kararı ile uygun görülmüştür.
Gereğini bilgilerinize rica ederim.

Prof. Dr. Ayhan GENÇLER
Başkan

Ek:Etik Kurul Kararı