A CORPUS-BASED ANALYSIS OF GRAMMATICAL COLLOCATIONS IN ARGUMENTATIVE ESSAYS OF TURKISH ELT STUDENTS

PhD Dissertation

Seda ÜNER

Eskişehir 2025

A CORPUS-BASED ANALYSIS OF GRAMMATICAL COLLOCATIONS IN ARGUMENTATIVE ESSAYS OF TURKISH ELT STUDENTS

SEDA ÜNER

PhD DISSERTATION

Department of Foreign Language Education

Program in English Language Teaching

Supervisor: Assist. Prof. Dr. Hülya İPEK

Anadolu University

Graduate School

May 2025

Eskişehir

JÜRİ VE ENSTİTÜ ONAYI

Seda Üner'in "A Corpus-Based Analysis of Grammatical Collocations In Argumentative Essays of Turkish ELT Students" başlıklı tezi 09/05/2025 tarihinde aşağıdaki jüri tarafından değerlendirilerek "Anadolu Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği'nin ilgili maddeleri uyarınca, Yabancı Diller Eğitimi Ana Bilim dalında İngilizce Öğretmenliği Programında Doktora tezi olarak kabul edilmiştir.

	Ünvanı Adı Soyadı	**İmza**
Üye (Danışman):	Dr. Öğr. Üyesi Hülya İPEK	**.....**
Üye:	Prof. Dr. Işıl KABAKÇI YURDAKUL	**.....**
Üye:	Prof. Dr. Aynur YÜREKLİ	**.....**
Üye:	Doç. Dr. Gonca SUBAŞI	**.....**
Üye:	Doç. Dr. Hüseyin KAFES	**.....**

Lisansüstü Eğitim Enstitüsü Müdürü

Prof.Dr.Nafiz Öncü CAN

ABSTRACT

A CORPUS-BASED ANALYSIS OF GRAMMATICAL COLLOCATIONS IN ARGUMENTATIVE ESSAYS OF TURKISH ELT STUDENTS

Seda ÜNER

Department of Foreign Language Education
Program in English Language Teaching
Graduate School of Anadolu University, May, 2025
Supervisor: Assist. Prof. Dr. Hülya İPEK

The significance of collocation is highlighted since it plays a crucial role in enhancing EFL/ESL learners' speaking, reading, listening, and writing skills, and facilitates more natural communication. It is often argued that non-native writers struggle to use grammatical collocations as effectively as native speakers, and even proficient learners who have achieved a significant degree of fluency in a foreign language may commit collocational errors. The present study aimed to reveal the use of grammatical collocations by Turkish learners of English through a corpus-based methodology. To this end, a learner corpus was compiled from argumentative essays written between 2009 and 2023 by prospective English teachers enrolled in the ELT Program at the Faculty of Education, Anadolu University. For comparison purposes, LOCNESS was used as the reference native corpus. The study examined learners' overuse and underuse of grammatical collocation patterns based on Benson et al.'s (2010) framework. Additionally, the study delved into the incorrect usage of grammatical collocations in the learner corpus and errors were analysed through an interpretative approach based on predefined error categories to identify potential sources of errors. The findings revealed that, despite some similarities in pattern usage, learners avoided certain patterns perceived as complex and tended to prefer simpler structures. The analysis of the sources of errors showed that intralingual errors outnumbered interlingual errors; however, L1 transfer errors were still notable. The study outlined methodological and theoretical implications for teaching grammatical collocations and emphasised the need to enhance their representation in textbooks and dictionaries based on the findings.

Keywords: Grammatical Collocations, Corpus, Argumentative Essays, Error Analysis, Turkish ELT Students

ÖZET

TÜRK İNGİLİZCE ÖĞRETMENLİĞİ ÖĞRENCİLERİNİN TARTIŞMACI YAZILARINDAKİ DİLBİLGİSEL EŞDİZİMLERİN DERLEM TABANLI ANALİZİ

Seda ÜNER

Yabancı Diller Eğitimi Anabilim Dalı
İngilizce Öğretmenliği Programı
Anadolu Üniversitesi, Lisansüstü Eğitim Enstitüsü, Mayıs 2025
Danışman: Dr. Öğr. Üyesi Hülya İPEK

Eşdizimlerin, EFL/ESL öğrencilerinin konuşma, okuma, dinleme ve yazma becerilerini geliştirmelerine katkı sağladığı ve doğal iletişimi desteklediği vurgulanmaktadır. Ana dili İngilizce olmayan yazarların, dilbilgisel eşdizimleri ana dili İngilizce olanlar kadar etkili bir şekilde kullanmakta zorlandıkları ve ileri düzeyde akıcılık kazanmış yabancı dil öğrencilerinin dahi eşdizim hataları yapabileceği belirtilmektedir. Bu çalışma, Türk İngilizce öğrencilerinin dilbilgisel eşdizim kullanımını, derlem temelli bir araştırma yöntemi doğrultusunda incelemeyi amaçlamıştır. Bu doğrultuda, Anadolu Üniversitesi Eğitim Fakültesi İngilizce Öğretmenliği Programı'na kayıtlı öğretmen adaylarının 2009-2023 yılları arasında yazdıkları tartışmacı kompozisyonlardan oluşan bir öğrenci derlemi oluşturulmuştur. Öğrenci derlemiyle karşılaştırma yapmak amacıyla, ana dil verisi olarak LOCNESS derlemi kullanılmıştır. Dilbilgisel eşdizim kalıplarının aşırı veya yetersiz kullanımı, Benson vd. (2010) tarafından önerilen sınıflandırma çerçevesinde analiz edilmiştir. Ayrıca, öğrenci derlemindeki hatalı dilbilgisel eşdizim kullanımları ayrıntılı olarak ele alınmış ve hataların potansiyel kaynaklarını belirlemek amacıyla önceden tanımlanmış kategorilere dayalı yorumlayıcı bir yaklaşımla analiz edilmiştir. Bulgular, kalıpların kullanımında benzerlikler görülse de, öğrencilerin karmaşık olarak kabul edilen belirli eşdizimleri kullanmaktan kaçındıklarını ve daha basit kalıpları kullanma eğiliminde olduklarını göstermiştir. Hata kaynaklarının analizi, içdilsel hataların, anadil aktarımı hatalardan daha fazla olduğunu, ancak anadil etkisinin yine de belirgin düzeyde olduğunu göstermiştir. Çalışma, dilbilgisel eşdizimlerin öğretimine yönelik yöntemsel ve kuramsal çıkarımları ortaya koymuş ve bulgular doğrultusunda, bu yapıların ders kitapları ve sözlüklerde daha kapsamlı şekilde yer alması gerektiğini vurgulamıştır.

Anahtar Sözcükler: Dilbilgisel Eşdizimler, Derlem, Tartışmacı Yazılar, Hata Analizi, Türk İngiliz Dili Eğitimi Öğrencileri

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Assist. Prof. Dr. Hülya İPEK, for her invaluable guidance, continuous support, and patience throughout my doctoral journey. Her insightful feedback, encouragement, and expertise have been instrumental in shaping this research.

I am also sincerely grateful to my thesis committee members, Prof. Dr. Işıl KABAKÇI YURDAKUL and Assoc. Prof. Dr. Gonca SUBAŞI for their constructive criticism, thought-provoking discussions, and unwavering support. Their expertise has greatly enriched my work and helped me refine my ideas. I would like to express my gratitude to the members of my defence committee, Prof. Dr. Aynur Yürekli and Assoc. Prof. Dr. Hüseyin Kafes for kindly giving their time, support, and direction through their wise remarks and helpful recommendations.

I would also like to thank fellow researcher Mehmet KAHRAMAN and my colleagues Berna DEMİREL and Jocelyn HERRERA for their technical assistance with this effort. My analyses were considerably more trustworthy and impartial because of their native speaker assessment of Turkish and English, as well as their background in language instruction.

A special thanks to my family, whose unwavering belief in me has been a constant source of strength. Their patience, love, and sacrifices have made this achievement possible.

Finally, I am deeply thankful to my friends, both inside and outside academia, who have offered emotional support, motivation, and much-needed moments of respite throughout this process.

This dissertation is the culmination of the collective support and encouragement I have received, and I am forever grateful to all those who have contributed to this journey.

Seda ÜNER

May 2025

09/05/2025

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Bu tezin bana ait, özgün bir çalışma olduğunu; çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı; bu çalışma kapsamında elde edilen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi; bu çalışmanın Anadolu Üniversitesi tarafından kullanılan "bilimsel intihal tespit programıyla tarandığını ve hiçbir şekilde "intihal içermediğini" beyan ederim. Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçları kabul ettiğimi bildiririm.

Seda ÜNER

09/05/2025

STATEMENT OF COMPLIANCE WITH ETHICAL PRINCIPLES AND RULES

I hereby truthfully declare that this thesis/dissertation is an original work prepared by me; that I have behaved in accordance with the scientific ethical principles and rules throughout the stages of preparation, data collection, analysis and presentation of my work; that I have cited the sources of all the data and information that could be obtained within the scope of this study, and included these sources in the references section; and that this study has been scanned for plagiarism with "Scientific Plagiarism Detection Program" used by Anadolu University, and that "it does not have any plagiarism" whatsoever. I also declare that if a case contrary to my declaration is detected in my work at any time, I hereby express my consent to all the ethical and legal consequences that are involved.

Seda ÜNER

09/05/2025

ÜRETKEN YAPAY ZEKÂ KULLANIM BEYANI

Bu tezi hazırlarken ChatGPT (ChatGPT, Gemini, DALL-E vb.) üretken yapay zekâ programlarından destek aldığımı beyan ederim. Tezin hazırlığı aşamasında üretken yapay zekâ programlarından bilimsel makale (örneğin çeviri, bilimsel makale vb.) desteği aldım. Üretken yapay zekâ programlarından aldığım bilgilerin doğruluğunu kontrol ettiğimi bildiririm. Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçları kabul ettiğimi bildiririm.

Seda ÜNER

09/05/2025

DECLARATION OF GENERATIVE ARTIFICIAL INTELLIGENCE USAGE

I declare that I received support from generative artificial intelligence programs (ChatGPT, Gemini, DALL-E, etc.) while preparing this thesis/dissertation. I have received support from generative artificial intelligence programs (e.g. language translation, access to scientific articles, etc.) during the preparation of the thesis/dissertation. I declare that I have checked the accuracy of the information I received from generative artificial intelligence programs. I also declare that, if a case contrary to my declaration is detected in my work at any time, I hereby express my consent to all the ethical and legal consequences that are involved.

Seda ÜNER

THE LIST OF CONTENTS

Page

THE LIST OF TABLES

Page

Page

Page

THE LIST OF FIGURES

Page

THE LIST OF ABBREVIATIONS

EA	: Error Analysis
ELT	: English Language Teaching
ELS	: English as a Second Language
EFL	: English as a Foreign Language
L1	: First Language
L2	: Second Language
LC	: Learner Corpus
LOCNESS	: Louvain Corpus of Native English Essays
SLA	: Second Language Acquisition
TL	: Target Language

CHAPTER 1 INTRODUCTION

1.1 Background of the Study

Learning the vocabulary of a language and using it correctly is an indispensable aspect of language learning. A crucial factor in the language learning process is the amount of vocabulary one possesses since vocabulary accounts for most of the meaning of a language (McCarthy, 1990); therefore, students who acquire an extensive vocabulary, in addition to possessing a comprehensive understanding of various types of vocabulary, are more likely to possess superior communication skills and a higher level of language proficiency (Vygotsky, 1986). According to Nation (2011), in both English as a Second Language (ESL) and English as a Foreign Language (EFL) context, vocabulary learning plays an important role in all language skills (i.e., listening, speaking, reading, and writing).

Many linguists agree that teaching and learning word combinations, particularly collocations, are essential for vocabulary development in second language classrooms. For instance, Lewis (2000), Nation (2001), and Bui (2021) concur that collocations play a crucial role in enabling EFL/ESL learners to use the target language with both accuracy and fluency, asserting that collocational competence is fundamental to achieving natural and proficient language use. As Nation (2001) puts it, “language knowledge is collocational knowledge” (2000, p. 522). The researcher further argues that enhancing one's ability to use collocations effectively is crucial for attaining fluency similar to that of native speakers because “all fluent and appropriate language use requires collocational knowledge” (p. 318).

Benson, Benson, and Ilson’s (1986) categorisation of collocations into two main types—grammatical and lexical—has significantly attracted researchers’ interest toward the study of collocations. Numerous researchers have compared the teaching of collocations and traditional vocabulary instruction, and they have also discovered compelling evidence that lexical retention is enhanced by the teaching of collocations (Balcı, 2006; Gilquin, 2007; Ghezelseflou and Seyedrezaei, 2015; Ördem and Paker, 2016; Bozoğlu, 2020).

Many researchers have emphasised the importance of collocation, as many studies have confirmed that collocation enables EFL /ESL learners to speak more fluently, improve their reading speed and listening comprehension, and write in a native-like manner (Brown, 1974; Pawley and Syder, 1983; Hill, 2000; Demir, 2018). According to Nesselhauf (2005), the speaker's ability to use language fluently is facilitated by the reduction of cognitive processing in the human brain when a large number of fixed units are present. Kamarudin et al. (2020) also underscored the significance of collocations, noting that an understanding of collocations is a reliable indicator of overall ESL proficiency.

Empirical studies of collocations have also demonstrated the fact that non-native speakers of English tend to use collocations inaccurately. Nesselhauf's (2003) research revealed that even advanced learners struggle to produce accurate collocations, aligning with Altenberg and Granger's (2001) findings. Consequently, Nesselhauf emphasises that "advanced learners' difficulties with collocations have not been investigated in much detail so far" (2003, p. 223). Additionally, Koç (2006) highlighted that Turkish students commonly experience significant difficulty in forming appropriate English collocations during their language learning process.

1.2 Statement of the Problem

The interaction of lexical items as a source of difficulty in vocabulary acquisition has been observed by researchers, who posit that learners' advancement should be measured by their mastery of these problematic combinations rather than their knowledge of individual words (Korosadowicz-Struzynska, 1980; Sadoughvanini et al., 2022). It is also highlighted that students encounter intralingual and interlingual issues when using collocations, and even proficient learners with significant fluency in a foreign language exhibit collocational inaccuracies (Du et al., 2022; Boonraksa and Naisena, 2022).

The literature reveals that non-native speakers have encountered the most frequent errors that pertain to the improper use of collocations across all error types (Gitsaki, 1999). The 'apparent rulelessness' of collocations has been noted by Laufer (1990) as one factor that interferes with foreign language vocabulary learning. The researcher further states that collocations are a crucial component of learners' vocabulary knowledge and acknowledges that difficulties may arise in their use of word combinations. Additionally,

the language-specific aspect of collocations may compel learners to create structures that are accepted in their first language but are unconventional or unattainable in their second language. The language chosen may interfere with the efficient use of the native language (Fan, 2009). This is because non-native speakers have less experience and often collocate words that sound foreign to native speakers (Ahmed, 2008).

In comparison with proficient writers, learners who compose texts in a language other than their first language may require a greater degree of awareness when employing these collocations (Groom, 2009). It is undeniable that these perplexing errors impede the transfer of knowledge to the reader, thereby restricting the author's persuasive capabilities (Sung, 2003). Nation (1990) asserted that low-proficiency learners "encode words in memory by sound and spelling, rather than by association with meaning" (p. 3). This is the reason why non-native writers are unable to employ collocations as effectively as native writers, as they are incapable of selecting words based on proper and specific word partnerships (Sung, 2003). Fox (1998) argues that the main problem in non-native academic writing, often marked by unusual expressions, lies in the improper use of collocations.

Based on previous studies in L2 writing, there are some common problems that L2 learners face, especially when writing an argumentative essay. According to Zhao (2017), argumentative writing is the most challenging writing genre that second language L2 learners face in universities since most of them do not have experience in writing academic texts in their first language (L1). L2 learners have difficulty in that genre due to the fact that argumentative writing requires persuading the audience and making them understand the other side of the argument by providing logical reasons to support a belief or idea (Wolfe et al., 2009). Writing an argumentative essay also includes critical thinking skills and organisational skills that involve planning the structure of an argument (Widyastuti, 2018). Therefore, the genre of argumentative writing is the one that requires the mastery of the language with the proper use of vocabulary and collocations. Although a number of language elements are required for a fluent writing process, both lexical and grammatical collocations are likely to occupy the first place in importance for non-natives. As McArthur puts it: Collocation knowledge is essential for fluency in writing, as miscollocations are considered “a major indicator of foreignness” (1992, p. 232). For this reason, studies comparing the production of native and non-native speakers are of

utmost importance in two ways: first, to identify the problems that non-native speakers face in the use of collocations, and second, to identify potential options and suggestions for the incorrect application of collocations.

A considerable number of studies investigating collocations have been conducted both within the Turkish context and across diverse international settings. In Türkiye, the existing body of research extensively explores a range of instructional methodologies, including explicit and implicit instruction, data-driven learning (DDL), and form-focused instruction (FFI). Furthermore, increasing attention has been given to the integration of digital technologies, particularly corpus-based instructional approaches and mobile-assisted language learning applications, in the teaching of collocations (e.g. Eker, 2001; Gençer, 2004; Balcı, 2006; Avcı, 2006; Şimşek, 2008; Akıncı, 2009; Öztuna, 2009; Kayıran, 2012; Öztuna, 2014; Aydoğan, 2016; Akkoyunlu, 2017; Özcan, 2017; Durmuş, 2019; Hamamcı, 2019; Sobucalı, 2019; Çepni, 2020; Aksungur, 2021; Türkmen, 2022; Genç, 2022). Various studies have also investigated the influence of collocation knowledge or training on various aspects of language proficiency (Demir, 2016; Qader, 2018; Bozoğlu, 2020). In addition, research on collocations in instructional materials, such as textbooks, has also received attention (Demir, 2022; Uçman, 2024). Within the Turkish EFL context, a substantial body of work has also been dedicated to investigating collocational errors (Shibliyev, 1993; Yılmaz, 2004; İlerten, 2010; Hama, 2010; Bıçkı, 2012; Üstünalp, 2013; Nişancı, 2014; Bağcı; 2014; Mutlu; 2015; Kahraman, 2021; Kayaer; 2023). However, among these error analysis studies, the focus has mostly been on lexical collocations, such as *verb-noun* patterns (Üstünalp, 2013; Bartan, 2019; Kahraman, 2021). While a few studies have addressed grammatical collocations, they tend to focus on limited structures (Yuvayapan and Yükselir, 2021; Dinçer et. al. 2022; Kayaer, 2023). For instance, the recent study by Kayaer (2023) investigated grammatical collocations with a particular emphasis on prepositions. Nonetheless, a comprehensive analysis of the full range of grammatical collocation patterns remains unexplored in the Turkish EFL context.

In light of the aforementioned issues, the present study focuses on the use of grammatical collocations in argumentative essays written by Turkish ELT students, comparing a learner corpus to a native speaker corpus. Conducting a thorough and authentic analysis of actual language use is essential to establishing a robust empirical foundation in

language research. In this regard, corpus-based methodologies serve as a powerful tool, providing contextually rich and accurate input for language learners. As Campoy et al. (2010) argue, corpus approaches allow researchers to gain a more realistic and comprehensive understanding of language by offering access to large-scale, real-world data across a variety of contexts.

One of the major advantages of this approach is its ability to move beyond subjective judgments or native speaker intuitions, which may not always reflect typical language use. Rather than relying on constructed or idealised examples that may lack authenticity and nuance, corpus data enable researchers to observe linguistic phenomena in naturally occurring discourse, spanning diverse genres, registers, and speakers. This access to authentic language use supports the formulation and empirical testing of hypotheses regarding specific linguistic features, thereby enhancing the reliability and validity of research findings (Klimova, 2014).

1.3 Purpose of the study

The primary objective of this research is to examine the utilisation of grammatical collocations in argumentative essays by Turkish ELT students whose proficiency levels range between B1 and B2. Using a native speaker corpus, the natural use of grammatical collocations in a learner corpus is investigated. Therefore, the study sought to examine the varieties and frequencies of grammatical collocations generated by Turkish ELT students and native speakers. In addition, it aims to identify the forms of grammatical collocational errors that exist within the learner corpus and the origins of these errors based on the categories of grammatical collocations.

Consistent with the above objectives, the following research questions will guide this study:

1. What are the types and frequencies of grammatical collocations appropriately produced by Turkish ELT students and native speakers?
2. What are the types of grammatical collocational errors within the learner corpus?
3. What are the sources of the errors within the learner corpus according to the types of grammatical collocations?

1.4 Significance of the Study

This study aims to develop a comprehensive understanding of Turkish ELT students' use of grammatical collocations in comparison to native-speaker productions in argumentative essays. To achieve this, the types and frequencies of grammatical collocations used by Turkish learners of English and native English speakers are systematically compared. This comparison is intended to reveal the similarities and differences in the use of grammatical collocations between the learner corpus and the native corpus. Additionally, the study seeks to gain a general understanding of whether Turkish learners of English use specific grammatical collocations as frequently as native English speakers. In the existing literature, research comparing native and learner corpora has predominantly focused on lexical collocations (e.g., Nesselhauf, 2005; Üstünalp, 2013; Kahraman, 2021), highlighting areas where language learners tend to diverge from native-like usage. While these studies offer valuable insights into lexical development, grammatical collocations have received comparatively less scholarly attention. By shifting the focus to grammatical collocations, the present study aims to address this gap and contribute to a more comprehensive understanding of learners' collocational competence. In doing so, it is hoped that the findings will inform which aspects of collocation instruction should be prioritised in English Language Teaching (ELT), thereby supporting more effective pedagogical practices and improving learner outcomes.

Second, the present study aims to conduct a detailed error analysis of grammatical collocation patterns found in the learner corpus. Specifically, it seeks to identify and categorise the types of grammatical collocational errors made by learners. While previous research has predominantly focused on lexical collocations or has limited its scope to specific grammatical collocation patterns (e.g. Yuvayapan and Yükselir, 2021; Dinçer et al., 2022; Kayaer, 2023), comprehensive analyses covering a wider range of grammatical structures remain scarce. Therefore, this study endeavours to address this gap by pinpointing the problematic areas in learners' use of grammatical collocations, thus contributing to a more nuanced understanding of learner difficulties and informing targeted instructional practices.

In addition to identifying error types, this study also aims to explore the underlying causes of these errors—an area often overlooked in existing research. Previous studies on

grammatical collocations tend to offer limited insight into the sources of errors, the gaps in learners' interlanguage, or the strategies they employ. In contrast, the present study places particular emphasis on uncovering the origins of errors through the lens of cognitive and communicative strategies. By doing so, it aims to offer a more comprehensive understanding of the cognitive and communicative strategies that lead to errors in learners' use of grammatical collocations. Ultimately, the findings are expected to inform the development of more effective instructional materials and pedagogical approaches for teaching collocations in the ELT context.

In conclusion, there is a substantial body of research that is dedicated to the examination of lexical collocations used by non-native Turkish speakers. However, it is worth noting that there is a notable scarcity of studies that specifically examine the grammatical collocations utilised by non-native speakers in Turkish contexts. This difference in research focus emphasises the need for further exploration of the grammatical aspects of collocations among non-native Turkish speakers, which could yield valuable insights into language acquisition and proficiency in this linguistic context. This study is unique in that it compares the grammatical collocations of non-native speakers based on a corpus of native speakers' argumentative writings in the Turkish setting. Additionally, it investigates the types and sources of errors in the use of grammatical collocations by non-native students. In this regard, the study also attempts to fill a gap in the existing literature. Ultimately, the objective of this research is to offer more precise recommendations for curriculum design, instructional strategies, and material development that will more effectively support learners' overall language proficiency and collocational competence.

CHAPTER 2 LITERATURE REVIEW

The first section (2.1) presents various definitions of collocation along with different approaches to defining the term, with specific attention to definitional frameworks discussed in subsection 2.1.1. Section 2.2 offers a chronological overview of collocation classifications within educational contexts. Section 2.3 focuses on the analysis of collocational errors, including discussions on error analysis (2.3.1), interlanguage theory (2.3.2), and the sources of collocational errors (2.3.3). Section 2.4 reviews empirical research on collocations, encompassing studies on the collocational competence of L2 learners (2.4.1), corpus-based investigations (2.4.2), research adopting Benson et al.'s (1986, 1997, 2010) classification of collocations (2.4.3), instructional approaches to collocation teaching (2.4.4), and academic theses on collocations conducted at Turkish universities (2.4.5).

2.1 Definitions of Collocation

Numerous definitions of collocation have been proposed in the literature. Firth (1957, p. 183), who first introduced the term in his work, defines collocation as "the company that words keep." According to McCarthy (1990), collocation constitutes a relationship which is "a marriage contract between words, and some words are more firmly married to each other than others" (p. 12). Sinclair (1991) adopts a broader perspective, defining collocations as instances where two or more words appear in close proximity within a text. Nattinger and DeCarrico (1992) define it as "strings of specific lexical items, such as rancid butter and curry flavour, that co-occur with a mutual expectancy greater than chance" (p. 36). According to Benson et al. (1997), collocation refers to words combined with certain other words or grammatical constructions. This means that there are words whose pairs can be put together to form natural semantic units. Therefore, it can be stated that expressive language can also be achieved through correctly made collocations. For Lewis (2000), collocations occur in a statistically significant way, and Hill (2000) points out that some collocations are predictable, e.g. the verb *foot*, which has a strong tendency to co-occur or collocate with the object noun *bill*, as in *foot the bill*. In addition, Nation (2001, p. 317) explains that the term collocation is used to refer to "a group of words that occur together either because they frequently occur together, such as *take a chance*, or because the meaning of the group is not evident from the meaning of the individual parts,

such as in *by the way* or *to take someone in*". According to *the Oxford Collocation Dictionary* (McIntosh et al., 2009), collocation combines words in a language to produce natural-sounding speech and writing. For example, incorrect combinations such as 'heavy wind' or 'strong rain' do not sound natural in English. Nesselhauf (2005) defines collocation as "a type of word combination in a particular grammatical pattern that refers both to an abstract language unit and to its instantiations in the text" (p. 25).

Martelli (2006), however, notes that "it is clear that not all scholars use the term to refer to the same phenomenon" (p. 1006). Some researchers emphasise frequency (e.g., Sinclair, 1991), while others prioritise the syntactic relationships among collocation components (e.g., Nesselhauf, 2003). Bahns (1993) similarly observes that "collocation is, unfortunately, a term that is used and understood in many different ways" (p. 57). Owing to this diversity in definitions and the broad scope of the concept, researchers have developed various approaches in an attempt to achieve a more thorough understanding and precise definition of collocations.

2.1.1 Approaches to define collocations

The complexity and variety of definitions are a result of the diverse methodologies that researchers have implemented when defining collocations. There are two main approaches to defining collocations: the phraseological approach and the frequency-based approach (Granger and Paquot, 2008).

The *phraseological approach* identifies collocations based on syntactic structure, semantic transparency, and substitutability (Nesselhauf, 2005; Laufer and Waldman, 2011) and provides a set of discrete criteria that can be used to categorise and analyse collocations. Within the phraseological approach, the *BBI Combinatory Dictionary of English* (Benson et al., 2010) divides collocations into lexical and grammatical collocations. Whereas lexical collocations consist of multi-word units with nouns, adjectives, verbs, and adverbs (e.g., verb-noun (*make dinner*), adjective-noun (*steady job*), verb-adverb (*speak softly*), grammatical collocations consist of a dominant word (e.g., noun, adjective, and verb) and a preposition or grammatical structure (e.g., noun-preposition (*reliance on*), adjective-preposition (*proud of*). The collocations included in the *BBI* are fixed, identifiable, and non-idiomatic, which has proven challenging for learners (Howarth, 1998). In line with this perspective, the phraseological approach

characterises collocations as non-idiomatic, but it also distinguishes collocations (fixed or identifiable phrases and constructions; e.g., *compose music, settle, adjust*) from free combinations (the elements that allow free substitutions according to grammatical rules; e.g., *want a car, at three of the clock, on break, in the library*. However, the distinction between collocation and a free combination is still based on the linguistic intuition of native speakers.

On the other hand, the frequency-based approach defines collocations as words frequently appearing together within a specific textual proximity. It differentiates between combinations that occur frequently—or more accurately, combinations appearing more often than random pairings would suggest—and those that do not (Nesselhauf, 2005, pp. 11-12). This perspective originates from Firth (1957) and was later expanded by Sinclair (1991), who describes collocations as two or more words occurring near each other within a text, with proximity determined by the lexical items involved (cited in Hong et al., 2011). To illustrate, Nesselhauf (2005, p. 12) provides an example where the word 'house' appears in the context, "He went back to the house. When he opened the door, the dog barked." In this sentence, the words 'went,' 'back,' 'to,' 'the,' 'when,' 'he,' and 'opened' are considered collocates of 'house.' The frequency-based approach thus does not treat collocations as fixed linguistic units but rather sees them as probabilistic phenomena. Consequently, from this viewpoint, there are essentially no impossible word combinations, though certain collocations naturally occur more frequently or commonly than others (Walker, 2011, p. 102).

In accordance with the aforementioned definitions and theoretical approaches, dictionaries have been created (Cowie, 1998). The focus of *the Oxford Dictionary of Current Idiomatic English – Vol. 1* (Cowie and Ronald, 1975) was to offer grammatical information for its entries. On the other hand, Volume 2 of the same dictionary (Cowie, Mackin, and McCaig, 1984) expanded its scope to encompass a broader range of idioms and collocations, drawing on phraseological models. Despite being more modern and corpus-based, *the Oxford Collocations Dictionary for Students of English* (McIntosh et al., 2009) remains consistent with phraseological principles, which assist students in selecting appropriate word combinations by analysing usage patterns. One of the most comprehensive collocation dictionaries, based on a phraseological framework*, is the BBI Combinatory Dictionary of English* (Benson et al., 2010). Given that the current study

adopts the collocation classification system outlined in the *BBI Combinatory Dictionary of English* (Benson et al., 2010), it follows a phraseological approach.

2.2 Classification of Collocations

Various categories of collocations have been adopted by different linguists (e.g. Hill, 2000; McCarthy and O'Dell, 2005); however, most of them classify collocations into two significant categories considering their syntactic features: one is called lexical collocations, and the other group is called grammatical collocations (Benson et al., 1997; Nesselhauf, 2005).

According to the most recent *The BBI Combinatory Dictionary of English* (Benson et al., 2010), lexical collocations are defined as phrases consisting of words such as nouns, adjectives, verbs, and adverbs, and they normally do not contain prepositions, infinitives, or clauses and do not have a dominant word. This dictionary categorises lexical collocations into seven main types, as seen in Table 2.1 below.

Table 2.1

The classification of English lexical collocations proposed by Benson et al. (2010) in The BBI Combinatory Dictionary of English

Type	Pattern	Example
L1	Verb + noun/pronoun or prep. phrase	*set a record*
L2	Verb + noun	*reject an appeal*
L3	Adjective +noun	*strong tea*
L4	Noun + verb	*blizzards rage*
L5	noun + of noun	*a pack of dogs*
L6	Adverb+ adjective	*deeply absorbed*
L7	Verb + adverb	*appreciate sincerely*

On the other hand, *The BBI Combinatory Dictionary of English* (Benson et al., 2010) defines *grammatical collocation* as a phrase consisting of a dominant word - noun, adjective/participle, verb - and a preposition or a grammatical structure such as an infinitive or a clause. According to this definition, in grammatical collocations such as *admiration for* (noun-preposition), *amazed at* (adjective-preposition), or *become something* (verb-complement), the dominant word is the noun or the adjective or the verb. Benson et al. (2010) categorise grammatical collocations into eight main types, and the eighth type (G8) has nineteen subcategories (G8(A)-G8(S)/(s)), as seen in Table 2.2 below.

Table 2.2

The classification of English grammatical collocations proposed by Benson et al. (2010) in The BBI Combinatory Dictionary of English

Type	Pattern	Example
G1	Noun + Preposition	*blockade against; apathy towards*
G2	Noun + to + Infinitive	*a pleasure to do it; the foresight to do it*
G3	Noun + that-clause	*He took an oath that he would do his duty.*
G4	Preposition + Noun	*by accident; in advance*
G5	Adjective + Preposition	*angry at; full of; fond of*
G6	Adjective + to + Infinitive	*ready to go; necessary to work*
G7	Adjective + that-clause	*She was afraid that she would fail the exam.*
G8(A)	Verb + D O + to + I O = V + I O + D O	*She sent the book to him. = She sent him the book.*
G8(B)	Verb + D O + to + I O	*They described the book to her.*
G8(C)	Verb + D O +for + I O = Verb + I O + D O	*She bought a shirt for her husband = She bought her husband a shirt.*
G8(D)/(d)	Verb + Preposition + Object / Verb + Object + Preposition + Object	*to act as; adhere to the plan* *We invited them to the meeting.*
G8(E)	Verb + to + Infinitive	*She continued to work.*
G8(F)	Verb + bare infinitive	*Mary had better go.*
G8(G)	Verb + V-ing	*They enjoy watching TV.*
G8(H)	Verb + Object +to Infinitive	*We forced them to leave.*
G8(I)	Verb + Object + Infinitive	*She heard them leave.*
G8(J)	Verb + Object + V-ing	*He felt his heart beating.*
G8(K)	Verb + a possessive +V-ing	*I cannot imagine their stealing apples.*
G8(L)	Verb + that-clause	*The doctor suggested that I take Vitamins.*
G8(M)	Verb + O + to be + Complement	*We consider her to be well-trained.*
G8(N)	Verb + O + Complement	*She dyed her hair red. She had her tonsils removed.*
G8(O)	Verb + O1 + O2	*The teacher asked the pupil a question.*
G8(P)	Verb + (O) + Adverbial	*He carried himself well.*
G8(Q)	Verb + (O) + wh-clause/ wh-phrase	*She asked why we had come.*
G8(R)	It + Verb + O + to Infinitive/ that-clause	*It surprised me to learn of her decision.* *It surprised me that our offer was rejected.*
G8(S)/(s)	Verb + Complement (Adj. or N)/ Verb + Complement (Adj.)	*He was a teacher. She was enthusiastic.* *The food tastes good.*

For classifying collocations, there exist other models proposed by researchers. For example, Hill (2000) agrees with Benson et al. and *the BBI* category on dividing lexical category into seven subcategories. However, the researcher notes that, comparatively, some elements are omitted or replaced with combinations involving more than two components (e.g., *adverb + verb*). Additionally, it is stated that certain collocations can be longer in structure; for example, an extended sequence such as *adverb + verb + adjective + noun + preposition + noun*, as in "*seriously affect the political situation in (Northern Ireland)*", is described as a semi-fixed expression.

Similar to the previous lexical models; *The Oxford Collocations Dictionary for Students of English* (McIntosh, 2009) lists a full range of collocations covering the seven

significant subcategories of lexical collocations, such as the model of Benson et al. (2010) and the BBI category. However, it adds subcategories for grammatical collocations that include prepositions (e.g., verb + preposition: *choose between two things*).

Additionally, McCarthy and O'Dell (2005) have presented six categories as examples (adjectives and nouns), (nouns and adverbs), (nouns and nouns), (verbs and expressions with prepositions), (verbs and adverbs), and (adverbs and adjectives). Moreover, they note that it can be challenging to distinguish collocations and compounds as essential and valuable components of vocabulary. Therefore, they include compounds as a category of collocations. Finally, they point out that idioms are groups of co-occurring words, but they are in a fixed order and have a meaning that cannot be guessed by knowing the meaning of the individual words. Consequently, they prefer to classify idioms separately from other types of word combinations.

In conclusion, this study addresses grammatical collocations based on the classification proposed by Benson et al. (2010). This classification was selected due to its ability to establish a clear and practical distinction between grammatical and lexical collocations, thereby providing a robust framework for systematic analysis. In addition, the taxonomy of Benson et al. is widely recognised and implemented in empirical research, as evidenced by a multitude of prior studies (e.g., Hsu, 2007; Wang and Shih, 2011). Adopting this well-established approach enables a more consistent and reliable examination of grammatical collocations within the current research context.

2.3 Analysis of Errors in Collocations

In addition to comparing learner and native corpora, this study also places significant emphasis on the analysis of grammatical collocation errors made by learners. Its theoretical framework is rooted in Error Analysis and Interlanguage Theory. Before outlining the key features of interlanguage, it is essential to acknowledge the foundational role of Error Analysis in understanding second language development.

2.3.1 Error Analysis

Error analysis (EA) has historically been a fundamental method in the investigation of language, with roots dating back to early language research (Sun and Shang, 2010). It has provided valuable insights into Second Language Acquisition (SLA) and English

Language Teaching (ELT) in several ways. Comparatively, error analysis seems to have taken the role of contrastive analysis, which extends SLA's behaviour theory. While error analysis, which is more closely related to the mentalism theory of SLA, views errors as normal for learners (Phoocharoensil, 2011).

As Ellis (1985) notes, error analysis can provide clues about the acquisition process. Additionally, analysing the linguistic errors made by students may serve two purposes: first, it provides information about students' current level of understanding; second, it helps educators create remedial activities that will help students perform better in their second language. According to Sari (2016), what still needs to be taught to a learner can be determined through error analysis. It gives the essential details regarding the areas in which he or she is deficient in competence.

Typically, EA researchers follow a common set of steps. They begin by collecting language data containing written or oral errors made by learners. These errors are then identified and classified into different types before being quantified. The next critical step is to determine the possible causes of these errors. If the purpose of the EA is pedagogical, the errors are also evaluated to assess their seriousness. This evaluation helps curriculum developers determine which types of errors should be emphasised in teaching materials (Ellis, 1985, 2008).

While EA has proven useful in language pedagogy, it has its limitations. One criticism is that EA commits the "comparative fallacy" by solely comparing learner language to target-language norms, overlooking the fact that learners develop their own unique rule systems, known as interlanguage, during the process of learning a second language (Selinker, 1992).

Furthermore, EA has been criticised for not providing a complete picture of learner language. Instead of solely investigating produced errors in second language learning, there is a call to examine the entirety of the linguistic system created by learners. It is argued that understanding the totality of learner production is crucial (Ellis, 2008). Dagneaux, Denness, and Granger (1998) propose limitations of error analysis as follows:

1. Heterogeneous Data: Traditional EA relies on diverse learner data, making it challenging to draw consistent conclusions.

2. *Fuzzy Categories:* Error categories in traditional EA can lack clarity and precision.

3. *Inability to Address Avoidance:* Traditional EA may not account for avoidance strategies employed by learners.

4. *Focus on What Learners Cannot Do:* It tends to concentrate on what learners cannot do rather than their correct TL usage.

5. *Static Perspective:* Traditional EA provides a static snapshot of L2 learning, lacking a dynamic view.

To address these limitations, a new type of EA emerged with the development of Learner Corpus (LC) research (Granger, 2009). LC-based error analysis shares similarities with traditional EA in terms of detecting and analysing errors but differs significantly in its approach. LC analysis offers several advantages, including systematic data collection, access to contextual information, and the ability to re-analyse and verify results. Moreover, LC research preserves text collections for further linguistic analyses (Nesselhauf, 2005).

Specialists in Second Language Acquisition (SLA) and English as a Foreign Language (EFL) recognise the importance of research on learner corpora (LC) as a valuable tool for gathering concrete evidence and gaining deeper insights into the characteristics and patterns of learners' interlanguage (Shih, 2000). This perspective acknowledges the importance of examining learners' linguistic development beyond simply identifying errors. Therefore, there has been a growing interest among language researchers in studying how learners develop linguistic competence in their own unique L2 system, known as interlanguage (Phoocharoensil, 2011).

2.3.2 Interlanguage Theory

Selinker (1972) defined interlanguage as the distinct linguistic system that is displayed when adult learners of second languages try to convey meaning in a language that they are still learning. This language system includes the lexical, pragmatic, and discourse levels of the interlanguage in addition to phonology, morphology, and syntax. There is a small variance in the definition of interlanguage. For instance, Corder (1971) defines interlanguage as transitional competence, which indicates the essential dynamism and

flux of the language learner's evolving system; however, Nemser (1971) refers to it as an approximative system, which is the state of linguistic competence of a learner at a specific moment in time. Interlanguage, as a linguistic system independent of both the native language and the target language, is universally accepted (Phoocharoensil, 2011).

In the study of second language acquisition, interlanguage is widely recognised as a viewpoint (Selinker, 1972). According to Ellis (1989), the interlanguage approach is centred on transitional linguistic systems that students build as they move closer to learning a second language. It recognises that mistakes are a normal component of learning and highlights the importance of systematicity in learners' language production. Researchers and educators can learn more about the fundamental cognitive processes involved in the acquisition of second languages by examining learners' interlanguage. According to Selinker (1972), interlanguage is the state that lies between a learner's native tongue and the target language. It is defined by systematic and rule-governed linguistic patterns. It acknowledges that language creation by learners is a unique system influenced by both their native language and the target language rather than just a reflection of one. The interlanguage approach holds that the interaction between input and feedback in the target linguistic and native language transfer accounts for errors made by learners as well as language progress. As a student is exposed to additional information and receives constructive criticism, their interlanguage system is dynamic and ever-evolving (Selinker, 1972). This viewpoint emphasises how crucial each learner's unique processes and methods are to pick up a second or foreign language.

Five basic cognitive processes were highlighted by Selinker (1972, 1992) as being particularly important in determining the formation of interlanguage, the transitional language system that second language learners establish during their language learning process. The following are these procedures:

a. Native Language Transfer: One important component influencing learners' interlanguage development is their native language (L1). Learning a second language (L2) can be aided or hindered by L1, depending on how much a student incorporates linguistic structures, elements, or habits from L1 into L2.

b. Overgeneralisation of Target Language (TL) rules: Learners frequently interpret L2 norms too broadly. This implies that individuals may appropriately apply a

general rule but fail to notice nuances or exceptions within it, which could result in mistakes in their interlanguage.

c. Transfer of Training: Students routinely put the rules they have studied in textbooks or from teachers into practice. The application may succeed if these rules coincide with the target language rules. On the other hand, students could create inaccurate interlanguage structures if the teaching is based on misconceptions or provides inaccurate information.

d. L2 Learning strategies: In order to master the target language, learners intentionally use a variety of tactics. These tactics could be mnemonic devices to help memorise vocabulary or other methods to improve language learning. Although these tactics can be useful, students run the risk of making mistakes if they lose their understanding of the ones they have developed.

e. L2 Communication strategies: Learners use communication methods when they are unable to communicate effectively because they do not have the linguistic components in the target language. These tactics could include meaning-conveying gestures, circumlocution, or other techniques. The language forms produced during these attempts might eventually become enduring components of the learners' interlanguage system.

To put it briefly, interlanguage functions as a transitional system between a learner's mother tongue and the target language. As Phoocharoensil (2011) explains, its defining characteristics include a systematic structure, the ability to integrate elements from both languages, its developmental nature, and the observable variation across different stages of progression. An understanding of interlanguage is essential for researchers and educators in the field of second language acquisition, as it offers valuable insights into the complex process of language learning and the strategies employed by learners.

2.3.3 Sources of Errors in Collocations

Several studies have demonstrated that EFL learners frequently encounter challenges when attempting to produce accurate collocations. These challenges lead to the emergence of distinct types of collocational errors, which can be attributed to various sources, including first language (L1) interference, overgeneralisation, paraphrasing, and target language (TL) interference.

Hong et al. (2011) constructed a framework that was inspired by Richards' (1974) and Tarone's (1981) frameworks, categorising these collocational errors based on cognitive and communicative strategies. The framework presented in Table 2.3 serves as a categorisation tool for identifying the sources of collocational errors.

Table 2.3

Framework for classifying the sources of collocational errors (Hong et al. (2011) adapted from Richards', 1974 and Tarone's, 1981)

Strategies	**Major categorisation of sources of errors**	**Sub-categorisation of sources of errors**
Cognitive strategies	Interlingual Transfer	a) L1 Transliteration / L1 literal translation b) Language switch
	Intralingual Transfer	a) False concept hypothesised b) Overgeneralisation c) Ignorance of rule restriction
Communicative strategies	Paraphrase	a) Approximation -semantic affinity -morphological and phonological affinities

The sources of errors were primarily classified into two fundamental concepts: cognitive and communicative strategies. Cognitive strategies were further divided into interlingual and intralingual transfer, encompassing five sub-categories in total. On the other hand, communicative strategies consist of a primary category, paraphrase, which was further subdivided into approximations with two components. In this study, additionally, a subcategory of intralingual transfer, specifically 'incomplete application of rules', was incorporated into this framework, drawing from Richards' (1974) study. In terms of interlingual transfer, instead of L1 transliteration and language switch, L1 interference was observed in a general term, as a transfer from the native language (L1) encompassing morphological, syntactic and lexical errors, in accordance with Tarone's (1981) study.

To better understand the underlying mechanisms behind these errors, it is essential to examine the two major categories of strategies involved in their production: cognitive strategies and communicative strategies.

a. Cognitive strategies

Cognitive strategies are methods and procedures people employ in their minds to help them learn, solve problems, and retain information. They facilitate the manipulation and organisation of new knowledge, making learners more effective (O'Malley and Chamot, 1990). A more comprehensive definition is provided by Sternberg (1986), who states that cognitive strategies are mental processes or techniques that promote the learning and memory of new information. Both intralingual and interlingual transfer are involved in cognitive strategies.

Interlingual transfer is the collective term for interference, language transfer, and cross-linguistic interference. These kinds of errors arise, according to Corder (1981), when a learner's ingrained habits, patterns, systems, or rules from their first language (L1) obstruct or hamper their ability to pick up patterns and rules in their second language (L2). Interference, also known as negative transfer, is defined by Lado (1964) as the detrimental effect of the learner's mother tongue (L1) on their proficiency in the target language (L2). A concise definition of interlingual errors is given by Chelli (2013), who states that language transfer is caused by the learner's first language. Transfer mistake is the root cause of interlingual errors, as stated by Allen and Corder (1974). According to Touchie (1986), L1 interference is the primary source of interlingual mistakes. Mahmoud (2019) points out that translation errors are the most difficult to explain since, in addition to the analyst's possible ignorance of the learners' native tongue, the identical mistake could also result via intralingual transfer. An EFL learner who speaks Arabic, for instance, produced *"I went to the home"*. The definite article and preposition were added incorrectly in English, possibly as a result of translation from Arabic, where the same parts are utilised in an identical structure.

Intralingual transfer is a term coined by Richards (1974:6), and it refers to errors made by learners that are based on generalisations made after only partial exposure to the target language rather than necessarily reflecting the structure of their mother tongue. These errors could be the result of learners still learning the target language and encountering difficulties when using it. According to Brown (2000), interlingual transfer—that is, interference from the native language—tends to predominate in the early phases of language acquisition. On the other hand, learners show more intralingual transfer or

generalisation inside the target language as they gradually pick up new language skills. Four categories, namely, *false concept hypothesised*, *overgeneralisation*, *ignorance of rule restrictions* and *incomplete application of rules,* are used by Richards (1974: 120) to categorise intralingual errors, each of which highlights the different ways that learners may make mistakes in the target language:

False Concept Hypothesised: This category includes misconceptions that arise from learners' incorrect conceptualisations of differences found in target language items. These mistakes are essentially the result of misinterpreting the subtleties and meanings of specific target language components (Al-Halalmeh, 2011). For instance, the use of de-lexical verbs like "make," "do," "go," "take," and so on interchangeably (e.g., "make homework") may lead to misconceptions about the subject matter (Ridha and Al-Riyahi, 2011).

Overgeneralisation: When learners construct abnormal language structures based on previous structures they have encountered in the target language, this is known as overgeneralisation. For instance, generalising the "-ed" past tense form or adding "s" to irregular plurals to make plurals may lead to overgeneralisation (Bıçkı, 2012).

Ignorance of Rule Restrictions: It is possible for learners to ignore some structural rules in the target language because they are unaware of them. Learners make this kind of error when they do not follow the rules of the language structures that are already in place. Analogical reasoning may lead to some rule restriction errors related to the rule-learning process. Ignorance of rule restriction is the term used to describe utilising the same linguistic elements of a specific structure that have been previously acquired on similar structures without considering their collocational and grammatical restrictions (Hong et al., 2011). For example, the use of improper propositions (e.g., look forward of) or the exclusion of propositions (e.g., travel the course) may result from a lack of comprehension of rule restriction (Li, 2005).

Incomplete application of rules: Such errors arise when learners are unable to fully apply the grammatical rules of the target language. This is frequently attributed to limitations in the linguistic input or contextual constraints that hinder the complete activation of those rules. These instances are indicative of an incomplete internalisation of the language system and reflect the learner’s inability to construct a fully developed interlanguage

framework necessary for accurate and consistent language production (Eun-pyo, 2002). For example, the learner partially applies the rules governing gerunds (e.g., insist on using) and, as observed by Selinker (1972), commits a developmental error.

b. Communicative strategies

The final classification concerns the communication strategies employed by students. Communicative strategies are defined as "a systematic attempt by the learner to express or decode meaning in the target language, in situations where the appropriate systematic target language rules have not been formed" (Tarone, Frauenfelder, and Selinker 1976; Tarone, Cohen, and Dumas 1976 as cited in Tarone, 1981:286). It encompasses the process of paraphrasing, which entails rephrasing an expression:

Paraphrasing: Within this category, there is the phenomenon of approximation, the use of a single target language vocabulary item or structure, wherein an incorrect vocabulary item or structure is used that nonetheless shares sufficient semantic properties with the intended item to satisfy the speaker (e.g. *pipe* for *water pipe*) (Tarone, 1981). In this framework, two types of approximation can be identified: morphological/phonological affinities and semantic affinities. These subcategories emphasise the morphological or phonological similarities to the corresponding form in L2, as well as the presence of related meanings. Illustrative examples include using '*noting someone*' instead of '*noticing them*' and '*entering the house'* instead of '*going inside*' (Ridha and Al-Riyahi, 2011).

These factors shed light on how learners construct their interlanguages, and interlanguage theory provides insights into learners' target language production in relation to these aspects. Moreover, these factors also explain the nature and origins of learner errors when analysing learner language. As Selinker (1992) noted, this system demonstrates the learner's developing comprehension and assimilation of the target language's rules and structures. According to Corder (1981), by analysing interlanguage for errors, teachers can gain clues about the extent of a learner's proficiency in the target language and identify areas where the learner may have deficiencies. Errors act as indicators of the learner's language system and areas that require improvement. Additionally, SLA researchers can utilise error analysis to delve into the process of L2 acquisition and gain insights into the strategies and methodologies learners employ during this process (Ellis,

2008). The analysis of errors provides researchers with a structured methodology for examining learner language. Learners themselves can use errors as a means to test their interlanguage hypotheses and develop their language skills. Making errors is an integral part of the learning process, and they serve as evidence of a learner's target language knowledge (Gass and Selinker, 2008). Finally, collocational problems, like other aspects of EFL learning, can be considered indicative of interlanguage development, as suggested by Higuchi (1999).

2.4 Empirical Studies on Collocations

A substantial body of research has examined collocations in second language acquisition, addressing collocation frequency, error analysis, instructional approaches, and L2 learners' collocational knowledge. The following sections provide a detailed review of studies focusing on collocational competence in L2 learners, corpus-based investigations, research grounded in Benson et al.'s (1986, 1997, 2010) classification, instructional studies on collocation teaching, and academic theses that explore collocations within the Turkish context.

2.4.1 Studies on the collocational competence of L2 learners

Numerous research studies have addressed learners' lexical and grammatical collocation knowledge. In one of the initial studies, Zhang (1993) conducted a study at Indiana University of Pennsylvania to investigate the writing of college students. The study entailed administering collocation assessments and collecting written work from a pair of participants: 30 native English speakers and 30 non-native speakers. The results suggested that native speakers outperformed their non-native counterparts, achieving significantly higher rates on both the collocation tasks and written assessments. This indicated that collocational proficiency plays a key role in enhancing fluency in written communication. In another preliminary study, Farghal and Obiedat (1995) investigated the collocational problems of students at Yarmouk University in Jordan. The researchers developed two versions of questionnaires: one in English with a "fill-in-the-blank" format and another in Arabic with twenty-two common collocations related to fundamental subjects, including food, colour, and weather. These questionnaires were given to English language instructors and senior and junior English majors. The findings demonstrated that there were notable deficiencies in collocations in both groups. The study also looked into

the individuals' task-solving skills. The strategies that were most frequently employed were transfer, avoidance, synonymy, and paraphrase. Expanding upon the study of collocational difficulties faced by EFL learners, Mahmoud (2005) investigated the accuracy of lexical and grammatical collocations in student writings, offering empirical support for previously identified patterns of unnatural word combinations. The researcher examined the use of collocations in 42 essays written by Arabic-speaking third-year university students majoring in English. The results revealed that approximately 64% of these collocations were incorrect, with the majority (80%) being lexical rather than grammatical collocations, and 61% of errors were due to negative transfer from Arabic.

Conversely, Shehata (2008) examined the factors that contribute to the enhancement of collocational knowledge. Specifically, the study examined two aspects: the learning environment as ESL/EFL and Arabic L1. The research was based on the analysis of two distinct participant groups: one consisting of 62 individuals studying English in an academic setting in Egypt, where English is taught as a foreign language, and another comprising 35 Arabic-speaking learners of English situated in a U.S. college environment, where English functions as a second language. In the ESL environment, participants were more proficient in mastering collocations, and their receptive perception of collocations was wider than their productive expertise. The study's findings indicated that the participants' learning environment and L1 (Arabic) influenced the acquisition of collocations. Further extending the analysis of collocational errors, Bahardoust and Moeini (2012) focused on the frequency and incorrect usage of lexical and grammatical collocations in the writing production of Iranian EFL learners across L1 and L2. The study involved 200 Persian EFL learners at the BA level who were taking paragraph writing and essay writing courses. The study categorised collocations into six groups: verb-noun, verb-adverb, be-adjective, adjective-noun, and verb-noun, and the analysis of the data included the evaluation of mid-term, final exam, and assignment papers of L2 learners using the chi-square test. Results showed that lexical collocations were more frequent than grammatical collocations, with verb-noun and adjective-noun collocations being the most common. Results also indicated that L1 interference had both positive and negative effects on the occurrence of collocations in the learners' writing.

The significance of negative transfer was also demonstrated in a study by Saud (2018), which aimed to identify the types of collocational errors made by undergraduate students.

50 undergraduate EFL students took part in the study. To determine the various kinds of errors in English collocation, a test with 24 multiple-choice questions was used to gather the data. The results showed that EFL students generally struggle with collocations, with *adjective + preposition* errors being the most common. *Verb + preposition* errors were the least frequent, followed by errors related to non-countable nouns. The main reason for these errors was the negative transfer from the native language to English and the lack of knowledge of collocation. Continuing the exploration of collocational errors, Sari (2017) specifically dealt with the students' errors in the grammatical collocation pattern *"verb + preposition"*. The sample consisted of 10 students, and the instrument of the study was a written test with 20 items. The responses in which verb-preposition combinations were used were analysed to find out why the students made errors. Consequently, the researcher found different reasons for the students' errors, namely, intralingual errors, interlingual errors, and carelessness. More specifically, Hudcovičová et al. (2021) conducted an empirical study of grammatical collocations of the type: *verb + for*. The object of analysis was the verbal-prepositional structures of pattern 1: *verb + prepositional phrase* in the object function and pattern 2: *verb + preposition + fix element*. The study was based on comparing English and Slovak sentences containing this specific verbal-prepositional structure. The research highlighted the complexity of translating verb-preposition collocations between English and Slovak, emphasising the need for context-aware translation strategies.

More recently, Boonraksa and Naisena (2022) expanded this line of research by examining lexical and grammatical collocation errors among EFL students at different proficiency levels, providing additional insights into how collocational accuracy develops over time. They compared the connection between first language (L1) and second language (L2) collocation inaccuracies in written assignments among high-, intermediate-, and low-level students. The research instrument consisted of 54 queries, separated into two sections, or collocation tests. The study used Lewis's (2000) concept of classifying collocations to investigate collocation errors made by EFL students. Initially, the investigation demonstrated that the grammatical collocation errors of the high-level EFL students were at a moderate level, while the medium and low-level EFL students were at a high level. All three groups of students were at a high level for lexical collocation errors. In addition, the collocation inaccuracies of the EFL students were the result of language

transfer from their first language (L1) to their second language (L2), synonyms, and a lack of collocation competence. Finally, the majority of the lexical collocation errors in all student groups were *adverb* + *adjective*. In the context of grammatical collocation, all groups were able to utilise *verb* + *preposition* more effectively than *noun* + *preposition*. *Adjective* + *preposition* was the least effective for students with high and medium proficiency, while *adjective* + *preposition* was more effective for students with low proficiency.

Sipayung and Saragih (2023) further explored collocation proficiency by examining the errors made by Indonesian EFL students. The subjects in this study were fifty students from a department of English instruction in Medan, Indonesia. A 50-item Simple Completion Test designed by the researchers served as the study tool. It evaluated students' knowledge of four types of lexical collocations: limited collocations, free combinations, figurative idioms, and pure idioms. The results indicated that these learners found pure idioms the most challenging, whereas free combinations presented the fewest difficulties. Regarding limited collocations and figurative idioms, student performance was about equal. The students' inconsistent answers frequently revealed a deficiency in their knowledge of English collocations. It was determined that the collocational errors were caused by inadequate first language transfer.

2.4.2 Corpus-based studies on the collocational competence of L2 learners

Corpus-based studies on the usage of collocations are also included in the literature, and their functions have attracted the interest of researchers who are interested in the challenges and errors that learners encounter when learning and generating a foreign language. Many scholars (e.g. Granger, 1998; Howarth, 1998; Nesselhauf, 2003, 2004, 2005) have analysed learner corpora and obtained valuable information about learners' collocation difficulties and errors. Their findings show that learners have a range of difficulties and problems with collocations. In general, they have indicated that learners' L1, the degree of similarity between L1 and TL, and the particular cognitive and communication skills they use can all impact the cause of collocation mistakes. In one of the most significant studies, Nesselhauf (2003) investigated the use of verb-noun collocations in written English by native Hebrew speakers at three different language levels. A learner corpus of argumentative and descriptive essays and a native speaker

corpus, LOCNESS, were used. In terms of the frequency of collocation use, learners were contrasted with native speakers, and regarding the correctness and frequency of collocations, learners were compared to other learners at varying levels of second language proficiency. The results indicated that learners at all three proficiency levels produced significantly fewer collocations compared to native speakers. An increase in the number of collocations was observed only at the advanced level, where specific interlanguage errors also began to emerge. In the same vein, Hsu (2007) investigated the use of collocations by non-native English speakers, with an emphasis on the writing of Taiwanese students. While Nesselhauf (2003) examined verb-noun collocation patterns across a range of Hebrew-speaking learners' proficiency levels, Hsu (2007) employed a web-based writing software to examine the frequency and diversity of collocations produced by Taiwanese students. The study's findings suggested a positive correlation between the writing grades of students and their collocation performances.

Building on this exploration of collocation use, Laufer and Waldman (2011) conducted an in-depth analysis of learners' language proficiency alongside their incorrect or excessively frequent use of collocations. To achieve this, they compared three different learner corpora with a native speaker corpus, which served as the reference point. The results revealed that all three non-native corpora exhibited lower performance in both the frequency and accuracy of collocation usage when compared to the native speaker data. Although the advanced learner corpus demonstrated a reduced number of errors relative to the other two non-native corpora, errors influenced by the learners' first language (L1) were consistently observed across all non-native datasets. Expanding on the investigation of collocational errors, Hong et al. (2011) examined verb-noun collocational errors in a Malaysian learner corpus, EMAS. The study involved 130 essays from Form Four Malay learners from three states. The analysis used the Interlanguage Theory and Error Analysis framework, Wordsmith Tools software, the Oxford Collocations Dictionary and the British National Corpus. The findings showed that preposition-related collocational errors were the most common, and intralingual transfer was the most prominent source of collocational errors.

Similarly, Phoocharoensil (2011) examined collocational errors, focusing on Thai EFL learners and the role of first language (L1) influence. While Hong et al. (2011) identified preposition-related errors as the most frequent among Malay learners, attributing them

largely to intralingual transfer, Phoocharoensil (2011) found that L1 interference was the primary cause of collocational errors among Thai learners. These errors included the addition or omission of prepositions, incorrect word choices, and redundant collocations. High-proficiency learners exhibited heavy reliance on their native language, while lower-proficiency learners resorted more frequently to their L1. Synonymy and overgeneralisation also contributed to deviations in collocations in the target language, emphasising the importance of L1 transfer and other strategies in these errors.

In terms of the further exploration of collocational errors, Hashemi et al. (2012) investigated collocational errors in English as a Foreign Language (EFL) college writing among 68 sophomores from Hamadan city. They used a modified framework and resources like the BBI Dictionary of English Word Combinations and the British National Corpus to analyse and provide corrections. The results of the study indicated that factors contributing to errors included native language interference, lack of understanding of collocational concepts, interlingual transfer, paraphrasing issues, and limited knowledge of collocations. Expanding the scope of collocational error analysis, Fadlilah (2016) specifically focused on grammatical collocation errors in the theses of English Department students, exploring their implications for teaching writing. The data obtained were then classified based on the English grammatical collocation patterns identified by the researcher. To simplify the data, the researcher selected and analysed the representative of each pattern from the total classified patterns of English grammatical collocations. The results of this study showed that students' mastery of the use of grammatical collocations in general and of collocation clusters in particular was weak. This finding was consistent with the general observation and claim of the research that word collocations are challenging even for advanced users of EFL.

Moreover, within the realm of corpus studies, researchers have explored both the frequency and diversity of the usage of lexical and grammatical collocations. In this context, studies have not only focused on identifying collocational errors but also on understanding the strength of collocational relationships to enhance language teaching. Hamer (2011) contributed to this line of research by investigating the collocation strength of 192 different *verb + preposition* combinations to determine which collocate strongly enough to warrant their teaching as single lexical units. The researcher examined a sub-corpus of the British National Corpus for Academic Writing and presented the verb-

preposition combinations that collocate most strongly. On the other hand, Güleç and Güleç (2015) dealt with genre comparison and investigated *verb + noun* lexical collocations in health, natural, and social sciences in written academic texts and analysed these lexical collocations using frequency and chi-square analysis. The aim of the study was to look for commonalities and differences between the verbs and their collocations. The results showed that there were more commonalities and relationships between the health sciences and the natural sciences, while the social sciences showed a significant difference compared to the other two. The study found 165 common verbs used in the three sciences. Among the 165 verbs analysed, twelve were recognised as representative prototypes of verb–noun collocations.

Building on previous corpus-based research, Xia et al. (2022) further examined collocation use by analysing the frequency, formulaicity, and diversity of lexical and grammatical collocations in argumentative essays from beginner to upper intermediate levels by using the Yonsei English Language Corpus. The results found that upper-intermediate L2 learners use more diverse and strong collocations than lower-intermediate learners. This suggests that certain patterns, such as *adverb + verb* and *verb + preposition*, can be used as solid markers to differentiate between novice and upper-intermediate L2 authors, as these collocations are more idiomatic and appropriate for academic writing. In the same year, Du et. al. (2022) studied learners' productive collocation competence using the British National Corpus and Python programming. They found that proficient learners use more complex and abstract collocations for communication tasks, while advanced learners use more difficult and longer nouns.

Complementing the focus on proficiency and collocation complexity, Dinçer et. al. (2022) investigated grammatical collocations in the written texts produced by A1 level Turkish foreign language learners of English. The study focused on the most frequent types of grammatical collocations produced by the students in the exams. The study revealed that students used 1791 tokens of nouns, 1186 adjectives, and 1043 verbs. Grammatical collocation types were used by 23.28% for *noun + preposition* formation and 27.32% for *verb + preposition* type, with the lowest usage being the *adjective + preposition* category (17.45%). Expanding on the analysis of grammatical collocations, Paris et. al. (2023) conducted a corpus-based study that looked into grammatical collocations aiming to investigate and evaluate the grammatical collocations used by Asian EFL students.

38,338 tokens from 20 academic writing papers written by Asian EFL learners were used to create an Asian learner corpus. The findings revealed that learners used similar prepositions in the top ten frequencies (of, to, in, for, on, as, by, from, with, and about) and the top three frequencies (of, to, and in) for each country. However, the order of prepositions for the fourth to tenth was different.

The corpus-based studies mentioned above examined the corpora of learners to explore the difficulties L2 learners face and the mistakes they make in learning and producing collocations. As Granger (2009) suggests, numerous studies indicate that applying corpus linguistic techniques to learner corpora helps us to identify the features that characterise learner language and to find out which features are shared by multiple learner groups and which are typical of a particular learner group, so that we can better understand how far learners' language is from the target language and what L2-specific patterns of overuse, underuse and misuse exist. Therefore, corpus-based studies have provided the field of ELT with valuable insights into L2 learners from different L1 backgrounds.

2.4.3 Studies on collocations based on Benson et al.'s (1986, 1997, 2010) classification

This section offers a comprehensive examination of numerous studies, utilising Benson et al.'s (1986, 1997, 2010) framework for collocation classification, as this investigation is based on their (2010) classification of collocations. Building on the investigation of grammatical collocations across different learner groups, Faghih and Mehdizadeh (2013) explored these patterns in both English and Azeri Turkish, focusing on the original English novella *Animal Farm* and its Turkish translation. The primary aim of their research was to uncover the similarities and differences in the grammatical collocation patterns between these two languages. They systematically identified and categorised grammatical collocations in the English novella, utilising Benson's (1986) classification of English grammatical collocations. Subsequently, these English collocations were compared to their corresponding Turkish counterparts. The analysis revealed significant differences in word order patterns between English and Turkish grammatical collocations, suggesting that Iranian Azeri Turkish speakers may face difficulties learning English grammatical collocations.

Using this taxonomy as a framework, numerous studies have investigated collocational errors. For instance, in their study, Namvar and Ibrahim (2014) investigated the influence

of learners' first language (L1) and cultural background on the production of collocations in the writing of postgraduate students using Benson et al.'s (1986) taxonomy. The findings indicated that learners encounter problems with both lexical and grammatical collocations in their texts and that the social and cultural context of the subjects have an impact on the generating of collocations. In the same year, based on the same taxonomy, Oktavera (2014) expanded on the exploration of collocational challenges by conducting qualitative narrative research on EFL students studying at the English Education Study Program in STAIN JuraiSiwo Metro. The researcher investigated the lexical and grammatical collocation performance of 10 EFL students in their argumentative writing. This research showed that 41 % of errors were found in lexical collocation (25%) and grammatical collocation (16 %) in the students' argumentative writing, which meant that students' collocation knowledge still needed to be improved.

Regarding the examination of collocation challenges among EFL learners, Alsulayyi (2015) attempted to investigate the production of English grammatical collocations among Saudi students majoring in English in KSA (Kingdom of Saudi Arabia) and in the UK. For this purpose, the researcher analysed the essays the participants wrote by using the classification of Benson et al. (1997). The highest number of errors in both groups was found for the grammatical collocation patterns *noun + preposition* and *adjective + preposition*. The results showed that Saudi EFL learners in the UK make fewer grammatical collocation errors than those learning English in KSA. The primary causes of these errors, as determined by this research, were L1 interference, avoidance, and a lack of understanding of grammatical collocations. Sari and Gulö (2019) also aimed to describe the types of grammatical collocation errors in essays produced by first-year Indonesian students majoring in English Literature at a private university. Collocations from fifteen essays were collected and analysed by the researchers. As a standard for the collocations to be analysed, they used the *Oxford Collocation Dictionary* (McIntosh, 2009) and the *BBI Combinatory Dictionary of English* (Benson et al., 1997). The results showed that students made grammatical collocation errors in their essays for 35 grammatical collocations. It was also found that *adjective + preposition* collocation was the most common error, indicating that students had difficulty using correct prepositions in collocations.

Bartan (2019) shifted the focus to lexical collocation errors in translation. By analysing Turkish-English translations of literary texts, the researcher compiled a written corpus to investigate the manner in which collocational errors are presented in their English translation texts (ETT), which were sourced from students enrolled in the Department of English Translation and Interpretation. To analyse the lexical collocation errors, the study employed Benson et al.'s classifications for lexical collocations, as outlined in 1997. Specifically, the research focused on the most common lexical collocation errors, particularly those involving *verb + noun* combinations. These errors were examined with regard to restrictions on collocability and the influence of the learners' native language (L1). The findings of the study revealed that the most prevalent type of error in ETT involved *verb + noun* collocations, followed by *adjective + noun* combinations. Furthermore, the research suggested that a higher degree of collocational restriction was associated with poorer production of collocations, underscoring the importance of understanding collocational constraints. Additionally, the study highlighted the significant role played by L1 influence in the generation of erroneous collocations among EFL/ESL learners and individuals undergoing translator education.

Aye (2021) focused specifically on investigating the grammatical collocation errors that postgraduate university students make when writing their essays in English. The investigation involved 45 first-year Myanmarese M.A. students, specialising in English. The errors were identified and analysed using the BBI Combinatory Dictionary of English by Benson et al. (2010). A total of 45 students in their first year of master's study were requested to compose a narrative essay. The grammatical collocation pattern G8 D / d (*verb + preposition + object (or) verb + object + preposition + object*) was the most frequently encountered error among the students. The researcher determined that the graduate students' errors in their writing were the result of their shortage of writing skills, knowledge of grammatical collocations, and the interference of their native language. Continuing the exploration of collocation use among EFL learners, Khorshid (2021) investigated the frequency of appropriate and inappropriate lexical and grammatical collocations among Iranian EFL students. 106 male and female students from Qazvin, Iran, ages 19 to 25, participated in the study. Benson's collocation approach and the Oxford Collocation Dictionary served as the foundation for the questionnaire that was used. The findings revealed that *noun + verb* collocations accounted for 73.6% of the

unsatisfactory lexical collocations. With 56.6% of them being incorrect, the *noun + preposition* type of collocations was the most commonly used. The study's findings indicated that learners were most prone to encountering challenges in generating appropriate collocations consisting of *noun + to infinitive* in terms of grammatical collocations and *verb + noun* regarding lexical collocations. The study sought to impact teacher preparation programs and curricula for Iranian EFL students.

In the Turkish context, Yuvayapan and Yükselir (2021) investigated the perceptions of lexical and grammatical collocations of a group of 19 EFL students who had taken a one-year English preparation course at a state university and examined their collocation errors and the reasons for them. Data were collected using an open-ended questionnaire, two focus group interviews, and descriptive and argumentative paragraphs written by students. The data from the questionnaire and focus group interviews were analysed using thematic analysis, and the students' writings were examined based on the same as the researchers aforementioned taxonomy of collocations proposed by Benson et al. (2010). The results indicated that lexical collocations were more frequently used, and the percentage of errors was higher in terms of lexical collocations. The participants' use of grammatical collocations was limited to five sub-categories: G1 (*noun + preposition*), G2 (*noun + to-infinitive*), G4 (*preposition + noun*), G5 (*adjective + preposition*) and G8 (*verb patterns*). Verb patterns were frequently used, and the G2 type (*noun + to-infinitive*) was also preferable among them. The researchers concluded that even though all participants were aware of the concept of collocation and its role in language learning, they made collocation errors when writing their texts, owing to the lack of automation of collocation, mostly caused by L1 interference.

Building on previous research into grammatical collocations, the distribution and accuracy of five categories of grammatical collocations G8E- G8I (Benson et al., 1986) were analysed by Sládková (2022) in CZEMATELC, an English language learner corpus consisting of 211,503 words. The corpus was derived from 1,841 English examination texts written by Czech secondary-school students. The results suggest that the A1-A2 CEFR level colligations are dominant, contingent upon a limited set of verb lemmas, and there is a significant amount of erroneous pattern variation. Additionally, there is a preference for patterns that are consistent with the most prevalent structures of their Czech counterparts. V inf was identified as the primary colligation in CZEMATELC,

comprising modal verbs and bare infinitives and accounting for approximately two-thirds of all patterns. A potential impact of the mother language was evident in the significant prevalence of verbs that largely conform to the most common pattern of their counterparts in the CNC (78.2%), as well as in certain mistake patterns, such as the V-pattern with centre modals and the erroneous V-inf pattern.

In the most recent study, Umair and Sarwat (2023) extended the investigation of collocation use by examining lexical and grammatical collocations in the narrative essay writing of undergraduate students. The researchers focused on examining the utilisation of lexical collocation, blunders, and syntactic collocation mistakes by the students of BS English at KFUEIT. They developed an essay-type test, and the collocation system and patterns by Benson and Ilson (1986) were applied to analyse the data to surface the collocation mistakes. The results showed that the frequency of lexical collocation errors was more significant than the frequency of grammatical collocation errors. Their research built on previous findings by focusing on how students at this academic level employ collocations in a specific genre of writing, offering fresh insights into their usage patterns and challenges.

2.4.4 Studies on instruction of collocations

Collocation instruction has been the subject of a significant amount of research, which has investigated the most effective methods for teaching lexical and grammatical collocations to language learners. In order to improve learners' collocational competence, research has investigated a variety of instructional methods, including explicit teaching, data-driven learning, corpus-based instruction, and communicative tasks.

Expanding the focus to instructional strategies, Fazlali and Shahini (2019) examined the impact of input enhancement and awareness raising on Iranian EFL learners' grammatical and lexical collocation knowledge. The study involved 60 participants, divided into a control and experimental group. Results showed that input enhancement did not significantly improve grammatical collocation knowledge, while awareness-raising instruction significantly improved lexical and grammatical collocation knowledge. The awareness-raising group surpassed the control group. Further examining the role of collocation instruction in language acquisition, the significance of collocations in English language acquisition for adult learners (students of military speciality) was investigated

by Sabanashvili and Garibashvili (2022). The researchers employed a multiple-choice questionnaire to determine that the instruction of collocation in EFL groups is one of the most straightforward methods for enhancing the communication skills of learners in both written and oral communication. Research into collocations also indicated that input enhancement techniques, such as providing passages with collocations highlighted, assisted in teaching collocation.

Corpus research has also informed English language teaching and learning by providing collocation lists and resources. Johns's (2002) explanation of "data-driven learning" (DDL) is where the term "corpus-based learning" originated, and it refers to language learners' use of corpora. The function of data-driven learning (DDL) in collocation instruction has been the subject of a significant amount of research. For instance, Chan and Liou (2005) examined the impact of five web-based practice units on English verb-noun collocations through the development of a web-based Chinese-English bilingual concordance program (Keyword Retrieval Program) on collocation learning. The study involved thirty-two college EFL students who completed a background questionnaire and an evaluation questionnaire, as well as a pre-test and two post-tests. The findings suggested that learners experienced substantial enhancements in their ability to learn collocations immediately following the online training, but they subsequently regressed. Nevertheless, the learners' initial performance was surpassed by the final performance. It was determined that learners with varying prior collocation knowledge and varieties of verb-noun collocations were not equally receptive to the practice effects.

Rahimi and Momeni (2012) conducted an additional experimental study to investigate the influence of teaching collocations on English language proficiency. The new words were taught to the control group in isolation for four months using conventional methods, including translation and definition. In the experimental group, vocabulary was taught through corpus-based activities and concordances, with an emphasis on teaching students collocations for a specific term. A fluency examination was administered to both groups subsequent to the experiment. According to the statistical analyses, the experimental group outperformed the control group post-test. This suggested that the instruction of collocations can enhance students' language proficiency. Li (2017) extended the exploration of corpus-based learning by comparing different instructional approaches. The researcher conducted a 15-week study in which 30 Chinese postgraduates were

compared to the effects of BNC (British National Corpus) and COCA (Corpus of Contemporary American English) based instruction and traditional teacher-led rule-based teaching (with access to dictionaries) in a linguistic course. The corpus group demonstrated significantly superior performance in the post-test (essay writing task) in terms of naturalness, the number of academic collocations, and fixed phraseological patterns.

A further study was conducted by Başal (2019) that compared collocation learning and retention scores between traditional exercises, such as multiple-choice and matching, and the learning of collocations using online collaborative collocation dictionary entries, concordance and the Web. The results indicated that the online resource produced significantly better learning and retention scores. Lay and Yavuz (2020) expanded the scope of research to include the reduction of errors in written language production. Their study aimed to investigate whether contextualised Data-Driven Learning (DDL) focused on addressing interlingual interference could help lower the frequency of errors in the writing of low-intermediate learners. The research involved 30 first-year Turkish university students enrolled in an English language program. Before implementing the DDL sessions, the researchers identified the ten most common error types attributed to L1 influence among Turkish learners of English. This listing was progressively facilitated by a literature review and the subsequent solicitation of the opinions of educators. In the subsequent phase, the participants were requested to compose essays in response to fundamental prompts, and the researchers anticipated that the predetermined target issues would be observed. The participants underwent an 11-week instructional session held in a computer lab, where each student had individual access to a computer. This phase was considered the pre-test stage. At the beginning of the treatment, students received training on using the TCSE corpus to locate and analyse the target linguistic structures. Upon completion of the treatment, a post-test was administered, and students' written compositions were assessed by comparing them to their pre-test performances, focusing on the specified language issues. The findings revealed that DDL-based instruction produced an average effect on learners' writing outcomes.

In another study, Santana (2022) addressed the impact of corpus research on collocations on the instruction and learning of the English language, particularly in English for specific purposes. The study concluded that corpus linguistics contributes knowledge of

collocations and corpus research informs English language teaching and learning. Building on the role of corpus research in English language teaching, Toomer et al. (2024) examined how Chinese speakers contextually learned English lexical (*verb* + *noun*) and grammatical (*preposition* + *noun*) collocations over the course of two days. By bolding or not bolding many occurrences of the target collocations in reading texts, the researchers were able to alter the learning circumstances. Collocational knowledge was assessed using a cued-recall (gap-fill) post-test. Responses, including grammatical collocations (but not lexical ones), were shown to be more accurate when they were bolded than when they were not. While reading, learners were likely to perceive *preposition* + *noun* collocations as more salient due to bolding. They concluded that explicitly focusing on the whole phrases in written input was likely to be beneficial for the contextual learning of L2 grammatical collocations.

Research on teaching English collocations has explored various aspects, including the proficiency of EFL learners in using collocations, the effectiveness of teaching collocations in EFL classes, the impact of input enhancement techniques on learning collocations, and the use of corpora in teaching collocations. Studies on contextualising proficiency show that EFL learners' proficiency in using English collocations varies and is influenced by multiple factors, including their linguistic background and the teaching methodology employed (Nattinger and DeCarrico, 1992). In addition, researchers have examined the influence of contextual learning, corrective feedback, and input frequency on the acquisition of collocations. Collectively, these studies underscore the significance of incorporating collocation instruction into language curricula to assist learners in the development of more natural and fluent language use.

2.4.5 Theses on Collocations in Türkiye

Theses on collocations in Türkiye can be broadly divided into four primary domains: error analysis on collocation use, the methodologies employed to teach collocations, the effects of collocation knowledge and training on learners' linguistic advancement, and collocations in textbooks.

Within the Turkish EFL context, there is a substantial number of theses that are dedicated to investigating errors in collocation usage, aiming to identify prevalent errors among language learners, examine their underlying causes, and offer insights into the acquisition

and use of collocations. Among these, Shibliyev (1993) investigated the utilisation of *verb-noun* collocations, specifically, '*make*' and '*do*' collocations among 36 beginner-level Turkish EFL students by means of a translation assignment and rating of acceptability. The test results showed that learners' unfamiliarity with the colloquial language, high restrictiveness of certain collocations, and prepositional collocations were the leading causes of errors. While Shibliyev (1993) investigated collocation errors among beginner-level Turkish EFL students, Yılmaz (2004) shifted the focus to advanced-level ELT students. This study investigated the characteristics of collocation errors at a higher proficiency level, offering additional insights into the obstacles that learners encounter when attempting to master collocational usage. A pre-test on collocations and idioms was administered to a control and experimental group. The experimental group was subjected to an eight-week treatment plan that was determined by their errors in the pre-test. Subsequently, both groups were administered a post-test. Error analysis indicated that collocation instruction significantly enhanced students' performance in the experimental group. This was demonstrated by the pre-test and post-test analyses.

In order to explore the extent to which second-language influence affects collocation use in the first language, İlerten (2010) examined the collocation knowledge of 22 Turkish immigrants in the USA, with 12 being first-generation immigrants and 10 being second-generation immigrants. Data was collected through the natural use of Turkish, biodata forms, multiple-choice tests, translation tasks, and cloze tests. The results were compared with sociolinguistic factors such as age of arrival, length of stay, and identity. The findings revealed that the immigrants' lexicon was influenced by the second language, with scores of the first-generation immigrants being close to those of the control group. In another study, Hama (2010) analysed errors in collocation formation and sought to discover error origins among upper-intermediate EFL learners, employing collocation exams and thinking protocols as data-gathering tools to get deeper insights into learners' difficulties. The learners' inaccurate performances were most significantly influenced by a lower incidence of specific collocations and L1 intrusion, as indicated by the evaluation of assessments and protocols.

Bıçkı (2012) expanded upon Hama's (2010) study of collocation errors in upper-intermediate EFL learners by analysing the structural specifics of erroneous productions

in adult advanced learners. This study sought to discover patterns in collocational errors in *verb-noun* combinations and investigate their underlying causes, providing greater insights into the linguistic obstacles faced by advanced learners. In the study, a total of 177 essays sourced from the ICLE-TRCU corpus, comprising writings produced by Turkish EFL learners. The researcher analysed 177 essays by using the ICLE-TRCU corpus, which contained written productions of Turkish learners of English. The results indicated that L1 generated a considerable impact on collocation mistakes. Notably, errors occurred more frequently with semi-restricted collocations and aspectual verbs. Üstünalp (2013) further explored errors and their sources, with a particular emphasis on *verb-noun* collocations. The researcher compiled a learner corpus from 434 essays of students in the ELT department. Then, the collocations were compared with the British National Corpus, and the Oxford Collocation Dictionary (McIntosh et al., 2009) was used to check the students' productions. The verb category and determiners were the most problematic components, and interference from L1 was the primary cause of error, in accordance with Bıçkı's (2012) findings.

In another study, Nişancı (2014) offered further insights into the specific challenges learners face in collocational usage. In a study involving 76 essays written by Turkish 11th-grade learners of English, the researcher focused on three primary sources of collocational errors: semantic opacity, limitations in permissible word combinations, and first language influence. The analysis and classification of the learners' incorrect usages revealed that the most challenging collocations were those with strict lexical constraints and low semantic transparency. In contrast, the impact of interference of L1 on collocational errors appeared to be relatively minimal. In her study, Bağcı (2014) broadened the scope of research on collocational competence by investigating how Turkish university-level EFL learners at varying proficiency levels use collocations. The study involved a comparative analysis of pre-intermediate and advanced learners, focusing on both receptive and productive dimensions of their use of lexical and grammatical collocations. Data were gathered through acceptability judgment tasks and cloze tests. The findings indicated that advanced learners outperformed their pre-intermediate counterparts in both understanding and producing collocations, highlighting the significant role of language proficiency in collocational development. Notably, while pre-intermediate learners showed no marked difference in their handling of lexical versus

grammatical collocations, advanced learners demonstrated greater accuracy with lexical collocations and committed fewer errors overall.

Mutlu (2015) transitioned the focus to the perceptions of Turkish EFL instructors and learners regarding the teaching and learning of collocations and analysed learners' errors. 32 high school instructors and 326 students from five public schools participated in the investigation. Results showed a positive correlation between comprehension of learners' vocabulary and the comprehension of collocations, but limited productive knowledge. The results of the study indicated that both learners and teachers accepted the significance of collocations, but learners felt teachers spent less time on instruction. Additionally, it was reported that the primary cause of error was the L1. In another study, Kahraman (2021) focused on a particular aspect of collocational usage by analysing the use of the high-frequency verbs 'make' and 'do' within verb-noun collocations in argumentative essays composed by ELT students. The learner corpus was compared to a native speaker corpus for reference, and patterns of overuse and underuse, grammatical and semantic structures, and instances of non-standard usage were analysed. The results indicated that learners considerably underutilised *make + noun* constructions in comparison to native speakers. Furthermore, the learner and native corpora exhibited both similarities and differences in terms of their semantic and grammatical characteristics. The error analysis also revealed that learners encountered difficulties in accurately generating verb-noun collocations that involved the verbs 'make' and 'do.' In the most recent study, Kayaer (2023) focused on the errors of ELT preparatory students in using English grammatical collocations, particularly prepositions. The study involved paragraphs, speaking performances, and a cloze exercise to identify errors. Based on the errors, a gap-fill task was developed and administered to identical individuals to ascertain the typical mistakes committed by the majority of the group. Subsequently, an error analysis was carried out on these typical errors with respect to generalisation within the target language and negative transfer from the mother tongue to the target language. The findings showed that participants make errors even when using grammatical collocations that are frequently used in the target language and that participants' performance is similarly impacted by generalisation and negative transfer.

In addition to error analysis, in the Turkish context, there is a considerable number of theses on collocations which encompasses studies investigating various instructional

methodologies, including explicit vs implicit teaching, data-driven learning (DDL), form-focused instruction (FFI), and the integration of technological resources such as mobile applications and corpus-based learning. These studies seek to identify the most efficacious ways for facilitating learners' acquisition and retention of collocations. In one of these studies, Eker (2001) examined the development of collocational competence in EFL learners. The study used compositions written during two-year writing courses as the primary data source. The data was analysed to categorise lexical collocation types and was examined using coding sheets. Key findings revealed that over time, fewer collocations were produced by learners, with Turkish learners showing better proficiency in certain subtypes of collocations. The influence of the learners' native language on lexical collocation production diminished as time progressed. Nevertheless, the frequency of transparent collocations reduced over time, suggesting that learners became more adept at using less straightforward or idiomatic collocations in their writing. While Eker (2001) investigated the longitudinal progression of collocational competence in students' written compositions, emphasising changes in collocation usage over time, Gençer (2004) employed a more controlled experimental methodology. The researcher presented a total of 36 upper-intermediate level EFL students at a state university with the same two texts to learn. The students in the experimental group were instructed to concentrate on the collocations in the texts. Subsequently, two assessments were administered to students to evaluate their productive and receptive abilities in relation to each text. The experimental group demonstrated superior performance on both assessments, which underscored the significance of explicitness in collocation instruction.

Expanding on research into effective vocabulary instruction, Balcı (2006) examined the impact of collocation instruction on 59 seventh-grade EFL students by comparing traditional methods of vocabulary instruction with collocation instruction. The learning and usage of new vocabulary were aided by the instruction of vocabulary by means of collocations, as evidenced by the comparison between the control and experimental groups. In another experimental study, Avcı (2006) explored the impact of incorporating explicit collocation instruction on the writing performance of pre-intermediate learners. Both the control and experimental groups participated in reading relevant texts and engaging in brainstorming sessions as a pre-writing activity. However, the experimental group received additional input in the form of suggested collocations related to the

vocabulary in the texts, and they were also provided with corrective feedback on collocational errors during the brainstorming phase. The results of evaluating the students' written outputs indicated that the experimental group did not outperform the control group in terms of overall writing quality. Nevertheless, their texts demonstrated a noticeably higher frequency of collocation use.

Şimşek (2008) further explored the effectiveness of different teaching approaches. The researcher compared teaching vocabulary through collocations to classical techniques. A total of 79 first-grade ELT students were allocated to two groups: the control group, which learned vocabulary through definitions, and the experimental group, which learned vocabulary through collocations. The experimental group exhibited higher retention rates, indicating the significance of collocations in vocabulary retention despite the absence of any substantial variation in vocabulary growth. In his study, Akıncı (2009) focused on lexical collocations, specifically verb-noun combinations. In a single group of advanced ELT students, the study contrasted the efficacy of three distinct treatments—data-driven learning (DDL), explicit instruction (EI), and combined method (C)—on the acquisition of verb-noun collocations. The researcher also incorporated the treatment preferences of participants in addition to performance evaluations. The results indicated that there was no significant difference in collocation performance between the three interventions. However, participants selected the DDL method based on the corpus guidance provided in the course of the research. Öztuna (2009) narrowed the scope of investigation to examine the use of *make* and *do* collocations among Turkish 7th-grade learners of English. The study involved three groups: a control group, a group exposed to Negative Evidence, and a group receiving Input Flood treatment. Post-test results revealed that both the Negative Evidence and Input Flood groups significantly outperformed the control group, suggesting the effectiveness of targeted input in improving collocational competence. The researcher suggested combining both explicit and implicit treatments for maximum benefit. Providing further insight into the function of collocation instruction in the development of language, Kayıran (2012) investigated the influence of collocation-focused instruction on the acquisition of vocabulary among ninth-grade students. The analysis of pre-test and post-test results revealed a significant improvement in their understanding of vocabulary by the study's conclusion, following a five-week intervention administered to a single cohort. Alongside the quantitative data, insights

obtained through a post-treatment survey suggested that students had also developed a more intense awareness of collocations.

In contrast to previous studies that primarily concentrated on traditional classroom-based methods, Öztuna (2014) introduced a technological dimension to the research on vocabulary acquisition by investigating the efficacy of collocation learning through an online platform. The researcher tested the computability of Laufer and Hulstijn's Involvement-Load Hypothesis with collocation instruction. The study, implemented in two elementary and pre-medium school groups, did not confirm the hypothesis due to a mismatch between theory and practice. The treatment was reportedly found to be excessively mechanical, confusing, and boring by some students. This was attributed to the mismatch between learner attitudes, cultural background, and educational policies. Öztuna (2014) presented a technical viewpoint on vocabulary acquisition by examining the efficacy of online platforms for collocation learning, whereas Aydoğan (2016) redirected attention to instructional methodologies. Aydoğan conducted a specific study on the impact of input-oriented and output-oriented tasks on the collocation improvement of upper-intermediate EFL students, providing more insights into effective pedagogical strategies. The evaluation took into account both passive and active retrieval. Participants were divided into input-oriented (control) and output-oriented (experimental) groups based on the instructional approach they received. Nevertheless, the pretest and post-test results revealed no significant difference between the two groups with respect to the two types of retrieval.

In addition to prior investigations into instructional strategies for the acquisition of collocational skills, Akkoyunlu (2017) embarked on a study that employed data-driven learning and teaching approaches to create a program designed for use in a translation course. The primary objective of this program was to enhance the awareness of lexical collocations among ELT students. To assess the effectiveness of the program, the study utilised a pre-experimental pretest-post-test design. The results showed that the program led to increased awareness of collocations and subsequently enhanced academic outcomes in the translation course. While Akkoyunlu (2017) concentrated on enhancing lexical collocation awareness through data-driven learning in a translation course, Özcan (2017) shifted the focus to the role of vocabulary knowledge and the integration of mobile applications in teaching collocations. The experimental study involved 54 students and

found a significant difference in academic achievement between the two groups, which indicated that students who used mobile apps during the learning process learned more meaningfully than those who learned in the classroom. The study also highlighted the potential of mobile applications in other educational settings, as they provide a permanent learning environment and can be used in various settings, including schools and multimedia materials.

Durmuş (2019) further refined the focus on collocation learning by conducting an experimental study to investigate the influence of two instructional approaches on the acquisition of *adjective-noun* collocations by secondary learners. The experimental group received implicit instruction, while the comparison group received explicit instruction. The study was conducted over a 10-week program, including pre-tests, post-tests, and delayed post-tests. It was revealed that the experimental group benefited from explicit collocation instruction, and the students retained most of the newly acquired knowledge. However, the Mann-Whitney U test showed no significant difference in collocation knowledge scores between the comparison and experimental groups. To offer a more comprehensive understanding of effective instructional approaches, Hamamcı (2019) expanded the investigation to first-year undergraduate students. This study employed a quasi-experimental design to investigate the influence of various teaching methods on the development of collocation knowledge. The dissertation concluded that both 'focus on form' and 'focus on forms' positively impact students' learning outcomes. While Hamamcı (2019) investigated the influence of various teaching methods on the collocation knowledge of first-year undergraduate students, Sobucalı (2019) conducted a comparative analysis of meaning-focused instruction (MFI) and form-focused instruction (FFI) on 19 university students with an elementary level of proficiency. Learners were separated into control, MFI, and FFI groups by the researcher. In accordance with the treatment of each group, 20 selected collocations were instructed. The experimental groups demonstrated a substantial improvement in comparison to the control group, as indicated by the results. Furthermore, the FFI group demonstrated superiority over the MFI group in the area of shape recognition.

Çepni (2020) investigated a data-driven methodology for collocation acquisition and examined the impact of two data-driven collocation learning methods (Corpus Consultancy and Practice on English-Turkish Parallel Texts) on participants' receptive

and productive collocational knowledge about form, usage, and meaning with a total number of 43 participants. The Corpus Group had training on utilising a corpus to ascertain and derive the meanings of the target collocations (10 adjective-noun and 10 verb-noun) via the COCA corpus. The Parallel Texts Group examined a limited corpus of English excerpts using target collocations sourced from the COCA corpus alongside their L1 translations. The participants' perceptions of both experimental procedures were gathered through a structured survey featuring open-ended questions. The results indicated that The Corpus Group surpassed the Control Group in both receptive and productive assessments, whilst the Parallel Texts Group's scores generally fell in between. Building on previous studies exploring various instructional approaches to collocation learning, Aksungur (2021) conducted an investigation into the efficacy of explicit teaching methods of L2 collocations that employed graphic organisers. The study involved 50 Turkish upper-intermediate students aged 18-19 who participated in English and reading/writing classes for 20 hours per week. It was found that students in the experimental group, taught with visual input, were better recallers and retained in memory longer.

Türkmen (2022) expanded upon Aksungur's (2021) discoveries regarding the advantages of explicit collocation instruction through graphic organisers by investigating the efficacy of Quizlet in EFL classes. This study investigated the effects of technology-enhanced vocabulary instruction on language acquisition and retention, involving 60 participants divided into three groups. The results were compared through pretest, post-test, and delayed post-tests, as well as a semi-structured interview. The study found that participants viewed Quizlet as an effective, practical, and enjoyable tool for vocabulary study. Moreover, it was proposed that Quizlet and collocations significantly contributed to vocabulary learning and retention. Following Türkmen's (2022) investigation into the efficacy of Quizlet for vocabulary acquisition, Genç (2022) investigated the efficacy of the Data-Driven Learning (DDL) method in the instruction of collocations in an experimental study. This research was specifically designed to enhance the academic writing skills of Turkish high school students by demonstrating the potential of corpus-based learning strategies to improve their collocational competence and writing proficiency. A mixed-methods research design was employed, with quantitative results indicating a significant difference between the control and experimental group.

Nevertheless, semi-structured interviews revealed negative student perceptions of the DDL approach, indicating a need for further research.

Apart from dissertations on error analysis and methods of teaching collocations, there are several studies examining the impact of collocation knowledge or training on multiple facets of language proficiency, such as vocabulary retention, writing abilities, speaking fluency, and overall communicative competence. Among these, Demir (2016) investigated the effect of the knowledge of collocations, lexical hedges, and lexical reinforcers in the articles from the ELT field by comparing Turkish authors with Anglophone authors. The researcher made a comparison between 100 academic articles from Turkish authors and the identical amount of articles from Anglophone authors. The Texts' analysis revealed that Anglophone authors employed a wider range and number of collocations, whereas Turkish authors overused specific collocations without demonstrating variation. This absence of diversity was concluded to impair the native-like fluency of their academic writing. While Demir (2016) examined the significance of collocational knowledge, lexical hedges, and lexical reinforcers in academic writing, the impact of collocation education on the speaking proficiency of learners was investigated by Qader (2018), hypothesising that knowledge of collocations could enhance fluency and communicative competence in EFL learners. 45 high school university pupils were assigned to either the control or experimental group. The results showed that instruction focused on collocations significantly improved students' speaking skills. The researcher concluded that mastery of collocations reduces speakers' anxiety and thus helps them speak more fluently. Expanding upon prior studies investigating the significance of collocations in language proficiency, Bozoğlu (2020) conducted a quasi-experimental study to investigate the effectiveness of teaching collocations on vocabulary knowledge and retention. The study involved 52 seventh-grade students, with 16 in the experimental group. The results showed that teaching collocations improved vocabulary learning and made it more enjoyable.

Finally, theses in the Turkish EFL context also include research on collocations in textbooks that aim to analyse their frequency, types, and pedagogical effectiveness in learning materials. For instance, Demir (2022) examined the pedagogical value of lexical collocations in three widely used intermediate-level EFL coursebooks designed for teaching general English in the Turkish context. The analysis focuses on the frequency

and types of lexical collocations assessed through their assigned usefulness scores in a corpus. The research assessed the quantity and types of collocations based on their usefulness ratings. The findings indicate that the New English File contained lexical collocations with the highest pedagogical value, making it the most beneficial for learners. In contrast, Language Leader and New Headway featured lexical collocations with relatively lower pedagogical value. In another study, Uçman (2024) expanded the analysis of collocations in English textbooks, focusing on the frequency of these collocations in 9th-grade textbooks published by the Ministry of National Education. The researcher compared these collocations with those in the British National Corpus, identifying overloaded or underused collocations. A descriptive analysis was conducted using questionnaires to assess the introduction and teaching of these collocations. Results showed that 66% of the textbooks were overused, while 34% were underused. The study also found that the textbooks lack sufficient support for effective collocation instruction and recommended revising, enhancing, and integrating collocation instruction into the curriculum in specific areas.

A review of relevant research in both Turkish and international contexts reveals that the majority of studies which evaluate the effectiveness of collocation training primarily concentrate on collocation instruction, corpus comparisons, or error analysis. Lexical collocations are the primary focus of most of these studies, which investigate the manner in which learners acquire and apply them in a variety of contexts. Although the majority of the current research has focused on lexical collocations, there are fewer studies that have comprehensively examined grammatical collocations and the specific errors that learners make in their usage. The bulk of existing research primarily deals with prepositional phrases, leaving other grammatical collocation patterns underexplored. This gap in the literature underscores the necessity of additional research into a variety of grammatical collocation patterns in order to gain a more comprehensive understanding of their usage in Turkish. In order to offer a more thorough comprehension of collocation acquisition, the present study endeavours to integrate corpus-based research with error analysis, recognising this gap. This study aims to identify and categorise errors in grammatical collocations, investigate their potential sources, and ascertain the extent to which learners' collocational usage differs from that of native speakers by comparing a learner corpus with a native speaker corpus. The objective of

this study is to provide valuable insights into the nature of grammatical collocational errors, the function of first-language interference, and the effectiveness of existing collocation instruction methods by combining corpus comparison and error analysis. The results are anticipated to make a significant contribution to the field by providing practical pedagogical implications for enhancing collocation training in foreign language education.

CHAPTER 3 METHODOLOGY

This chapter provides an overview of the research design (3.1), the corpora employed in the study (3.2), the learner corpus (3.2.1) and the native corpus (3.2.2). Section 3.3. addresses the procedure which gives details on syntactic patterns included in the study (3.3.1), identification of the grammatical collocation patterns (3.3.2) and data analysis (3.3.3).

3.1 Research Design

The current study adopts a corpus-based methodology and is primarily descriptive in nature. As Kennedy (2003) noted, the use of corpus-based methods is a common and reliable approach to accurately describe language structure and acquisition. Rather than attempting to investigate previously unknown linguistic elements, corpus-based research focuses on analysing how established linguistic elements function within consistent patterns of real-world language use, as Biber (2010) points out. This study seeks to conduct an in-depth examination of grammatical collocations—an established aspect of language—by analysing their occurrence in the authentic language production of both English learners and native speakers.

To assess whether learners either overuse or underuse these specific combinations, the study uses a chi-square test. The chi-square test is a statistical method used to assess the association between two categorical variables and to evaluate whether the discrepancies between observed and expected frequencies are statistically meaningful or simply the result of random variation, as Field (2005) explains.

Additionally, the study utilises descriptive analyses, including frequencies and percentages, to elucidate the distinctions between the two corpora. Furthermore, any instances of incorrect productions in the learner corpus are examined through an interpretative approach, taking into account predefined categorisations.

3.2 Research Corpora

Two distinct collections of texts were employed in this study. The initial corpus comprises essays produced by Turkish learners of English at the B1–B2 proficiency level. The second corpus, serving as a reference, comprises essays authored by native English

speakers. Notably, both corpora consist of essays generated by learners, specifically in response to specific argumentative prompts.

The idea of corpus representativeness has been crucial in the field of corpus research (Biber, 1993; Biber, Conrad, and Reppen, 1998). The *size* and *composition* are two important factors in determining representativeness when comparing corpora. The term *size* relates to the necessity for each corpus being compared to be large enough to reflect the frequency and range of linguistic phenomena reliably. In contrast, *composition* involves the intentional selection of corpus content to include a variety of registers that reflect real-world language use (Biber, 2010). Therefore, both *size* and *composition* must be carefully considered when constructing or comparing corpora, as they directly affect the accuracy and generalizability of any linguistic analysis.

It is essential to acknowledge that, irrespective of *size*, corpora can only capture a limited portion of a language. As a result, the key consideration becomes determining the number of words required to achieve a descriptively adequate representation. In practical terms, especially when working with learner corpora, the datasets are often relatively small in scale. Nelson (2010) stated that the necessity of smaller corpora, which focus on specific aspects of the language, has been acknowledged, particularly in the context of language classroom instruction and utilisation. Kennedy (1998, as cited in Nelson 2010, p. 68) emphasised that researchers should "bear in mind that the quality of the data they work with is at least as important as the size." Davies (2015) also pointed out that although large corpora are generally preferred, ensuring accurate annotation becomes increasingly difficult as corpus size expands. In other words, the larger the dataset, the higher the risk of inaccuracies in areas such as tagging words with the correct part of speech (lexical-level annotation) and categorising texts according to predefined parameters (text-level annotation).

Given the insights from the literature, the current study aimed to achieve the best possible level of representativeness for the corpora used. To ensure the reliability of the comparison, several aspects were taken into account, such as the alignment of essay topics and the equivalence in corpus size. A comprehensive overview of each corpus is presented below.

3.2.1 The learner corpus

The learner corpus under investigation comprises argumentative essays authored by first-year students enrolled in the Department of English Language Teaching at Anadolu University during the period from 2009 to 2023. These students include those who completed the English preparatory year before entering the department and successfully passed the end-of-term proficiency exam, as well as those who passed the initial proficiency exam. These students are considered to possess a level of English proficiency between B1 and B2 (intermediate to upper-intermediate). The study incorporates a total of 206 essays, collectively containing 172,509 words. The essays were all obtained from the Turnitin system, where students submitted their assignments in their academic writing courses during the designated timeframe. The essay topics encompassed a range of themes, including education, technology, social life, health, media, and art.

3.2.2 The native corpus

To enable meaningful comparisons, this study incorporates the Louvain Corpus of Native English Essays (LOCNESS), compiled by Granger (1998). LOCNESS encompasses a total of 168,593 words drawn from argumentative essays written by university students from both Britain and the United States. It is noteworthy that all the essays included in LOCNESS were authored by young adult native English speakers. 33 of them were written by British university students, and 176 of them were written by American university students. These essays serve as argumentative responses to various prompts, which include topics such as homelessness, capital punishment, drug abuse, gun control, recycling, football, and the values and consequences of school interaction.

The essay subjects were education, technology, social life, health, media, art, sports, environment, and politics, despite the fact that the clue words demonstrate a wide range of variability. Comprehensive statistical information pertaining to both corpora can be found in Table 3.1.

Table 3.1

The corpora in the study

	Learner Data	**Native data (LOCNESS)**
Subjects	Turkish university students at B1-B2 level of proficiency	American and British university students who are native speakers of English
Essay Genre	Argumentative	Argumentative
Number of Essays	206	209
Total Corpus Size	172,509 words	168,593 words
Average Essay Length	910 words	850 words
Essay Topics	Education Technology Social life Health Media Art	Education Technology Social life Health Media Art Sports Environment Politics Gender issues

3.3 Procedure

This section begins with an introduction to the selected syntactic patterns that serve as the search criteria. Following that, an overview of the process of identifying the grammatical collocation patterns and a detailed exposition of the data analysis procedure are presented.

3.3.1 Syntactic patterns included

In this study, all the grammatical collocation patterns that correspond with Benson et al.'s (2010) classification are included. There are 8 main patterns of grammatical collocations: G1, G2, G3, G4, G5, G6, G7, and G8. G8 collocations consist of nineteen English verb patterns, designated by the capital letters A to S (G8(A)-G8(S)). Consequently, a total of 26 patterns were subjected to the search criteria, as illustrated in Table 3.2.

Table 3.2

The grammatical collocational patterns included (Benson et al.,2010).

Type	Pattern	Example
1. G1	Noun + Preposition	*blockade against; apathy towards*
2. G2	Noun + to + Infinitive	*a pleasure to do it; the foresight to do it*
3. G3	Noun + that-clause	*He took an oath that he would do his duty.*
4. G4	Preposition + Noun	*by accident; in advance*
5. G5	Adjective + Preposition	*angry at; full of; fond of*
6. G6	Adjective + to + Infinitive	*ready to go; necessary to work*
7. G7	Adjective + that-clause	*She was afraid that she would fail the exam.*
8. G8(A)	Verb + D O + to + I O = V + I O + D O	*She sent the book to him. = She sent him the book.*
9. G8(B)	Verb + D O + to + I O	*They described the book to her.*
10. G8(C)	Verb + D O +for + I O = Verb + I O + D O	*She bought a shirt for her husband. =She bought her husband a shirt.*
11. G8(D) /(d)	Verb + Preposition + Object / Verb + Object + Preposition + Object	*to act as; adhere to the plan We invited them to the meeting.*
12. G8(E)	Verb + to + Infinitive	*She continued to work.*
13. G8(F)	Verb + bare infinitive	*Mary had better go.*
14. G8(G)	Verb + V-ing	*They enjoy watching TV.*
15. G8(H)	Verb + Object + to Infinitive	*We forced them to leave.*
16. G8(I)	Verb + Object + Infinitive	*She heard them leave.*
17. G8(J)	Verb + Object + V-ing	*He felt his heart beating.*
18. G8(K)	Verb + a possessive +V-ing	*I cannot imagine their stealing apples.*
19. G8(L)	Verb + that-clause	*The doctor suggested that I take Vitamins.*
20. G8(M)	Verb + Object + to be + Complement	*We consider her to be well-trained.*
21. G8(N)	Verb + Object + Complement	*She dyed her hair red. She had her tonsils removed.*
22. G8(O)	Verb + Object 1 + Object 2	*The teacher asked the pupil a question.*
23. G8(P)	Verb + (Object) + Adverbial	*He carried himself well.*
24. G8(Q)	Verb + (Object) + wh-clause/ wh-phrase	*She asked why we had come.*
25. G8(R)	It + Verb + Object + to Infinitive / that-clause	*It surprised me to learn of her decision. It surprised me that our offer was rejected.*
26. G8(S)/(s)	Verb + Complement (Adj. or N)/ Verb + Complement (Adj.)	*He was a teacher. She was enthusiastic. The food tastes good.*

According to the *BBI Collocation Dictionary* (Benson et al., 2010), which serves as a reference for this study, grammatical collocation patterns are described as follows:

G1 collocations consist of *noun + preposition* combinations such as *blockade against; apathy towards*. *Noun + of* combinations are not included in this category as they represent the notion of the concepts of 'direct object', 'subject', or 'possession'. *Noun + by* combinations and free combinations, such as *apathy among (the members of our party), apathy in (France)* are not included either.

G2 collocations are comprised of *nouns* followed by *to + infinitive*. Infinitives expressing purpose (*in order to*); or phrases like *a procedure to follow, a book to read*, are not included in this category.

G3 collocations are *nouns* that can be followed by a *that clause*: e.g. *we reached an agreement that she would represent us in court*. In this category, nouns followed by relative clauses introduced by that, i.e., when that can be replaced by which (e.g. we reached an agreement that would go into effect in a month), are not included.

G4 collocations consist of *preposition + noun* combinations. Examples are: *by accident, in advance, to somebody's advantage, on somebody's advice, in agony, on (the) alert, etc*.

G5 collocations contain *adjective + preposition* combinations in the predicate or as set-off attributives (verbless clauses): *they were angry at everyone — angry at everyone*. Derived prepositions such as *concerning, regarding, in regard to, with regard to,* are not included in this category. In general, past participles followed by the preposition *by* are not included (*abandoned by, absolved by, etc.*) either.

G6 collocations consist of predicate *adjectives* and a following *to + infinitive*. Examples are, *it was necessary to work*, *she (the girl) is ready to go; it (the machine) was designed to operate at high altitudes*.

G7 *adjectives* (many of which are also in G6) can be followed by a *that clause*: *she was afraid that she would fail the examination; it was nice that he was able to come home for the holidays.*

G8 collocations consist of 19 English verb patterns, designated by the capital letters A to S. Each verb pattern is described as follows:

In G8(A) verbs, the dative movement transformation is permissible, which enables the relocation of an indirect object to a position preceding the direct object. The word "to" is omitted when both objects are nouns and when the direct object is a noun: *he sent the book to his brother — he sent his brother the book.*

G8(B) verbs are transitive; when they have an indirect object, they do not allow the dative movement transformation, i.e., the shift of the indirect object (usu. animate) to a position before the direct object with the deletion of to: *they described the book to her*, but not *they described her the book.*

In G8(C), the dative movement transformation is enabled by transitive verbs that are used with the preposition for. This transformation allows the indirect object (usu.animate) to be shifted to a position before the direct object and the deletion of "for": *she bought a shirt for her husband — she bought her husband a shirt*.

In G8(D/d), the verb forms a collocation with a specific *preposition (+ object)*. Collocations consisting of a verb + as (+ object) are included: *to act as, to interpret as, to serve as, to treat as, etc*. When a verb (or a certain sense of a verb) is normally followed by a prepositional phrase, its pattern is designated by the small letter '*d'*.

In G8(E), verbs are followed by *to + infinitive*. Examples of this construction are: *they began to speak, she continued to write*, etc. Verbs are excluded from this category if they are typically employed in purposeful phrases: *they were drilling (in order) to improve their pronunciation*.

G8(F) includes the small number of verbs that are followed by an infinitive without to: *we must work*, *he had better (would rather) go*.

In G8(G), verbs are followed by a second verb in *-ing*: *they enjoy watching television, he kept talking, we miss going to work every day, the house needs painting, she quit smoking*, etc.

In G8(H), transitive verbs are followed by an object and *to + infinitive*: *she asked me to come; they challenged us to fight; we forced them to leave; he invited me to participate*, etc.

G8(I) includes transitive verbs followed by a direct object and an infinitive without to: *she heard them leave, we let the children go to the park*, etc.

G8(J) consists of verbs followed by an object and a verb form in *-ing*: *I caught them stealing apples*, *we found the children sleeping on the floor*, etc.
In G8(K), verbs can be followed by a possessive (pronoun or noun) and a gerund, i.e., a verbal noun: *please excuse my waking you so early, this fact justifies Bob's coming late*, etc.

In G8(L), verbs can be followed by a noun clause beginning with the conjunction that: *they admitted that they were wrong, she believed that her sister would come*. In colloquial English *that* may be omitted: *they admitted they were wrong*, etc.

In G8(M), a direct object, the infinitive "to be," and either an adjective, a past participle, or a noun/pronoun can follow transitive verbs. These are some examples of this construction: *we consider her to be very capable — we consider her to be well trained — we consider her to be a competent engineer.*

G8(N) includes transitive verbs that can be followed by a direct object and an adjective or a past participle or a noun/pronoun: *she dyed her hair red, we found them interesting, he made his meaning clear, the police set the prisoner free*.

In G8(O), transitive verbs are capable of accepting two objects, neither of which is typically appropriate for use in a prepositional phrase with to or for. Sentences that contain such double objects include: *the teacher asked the pupil a question, we bet her ten pounds, the police fined him fifty dollars, God will forgive them their sins, she tipped the waiter five dollars*, etc.

G8(P) consists of intransitive, reflexive, and transitive verbs followed by an adverbial. The adverbial may be an adverb, a prepositional phrase, a noun phrase, or a clause: *he carried himself well; or: he carried himself with dignity*.

In G8(Q), verbs can be followed by an interrogative word: *how, what, when, where, which, who, why* and *whether/if*: *he wants what I want.*

In G8(R), the dummy it precedes transitive verbs, which are frequently used to convey emotion. They are followed by either the *to + infinitive*, the *that + clause*, or both. Examples are: *it behooves/behoves you to study more; it puzzled me that they never answered the telephone*.

In G(S/s), Only a small number of intransitive verbs are followed by a predicate noun or a predicate adjective: *she became an engineer*; *she was enthusiastic*. The verb "make," when used intransitively, is appropriate for this context: *he'll make a good teacher*. A predicate adjective is the sole suffix that can follow a string of intransitive verbs; these

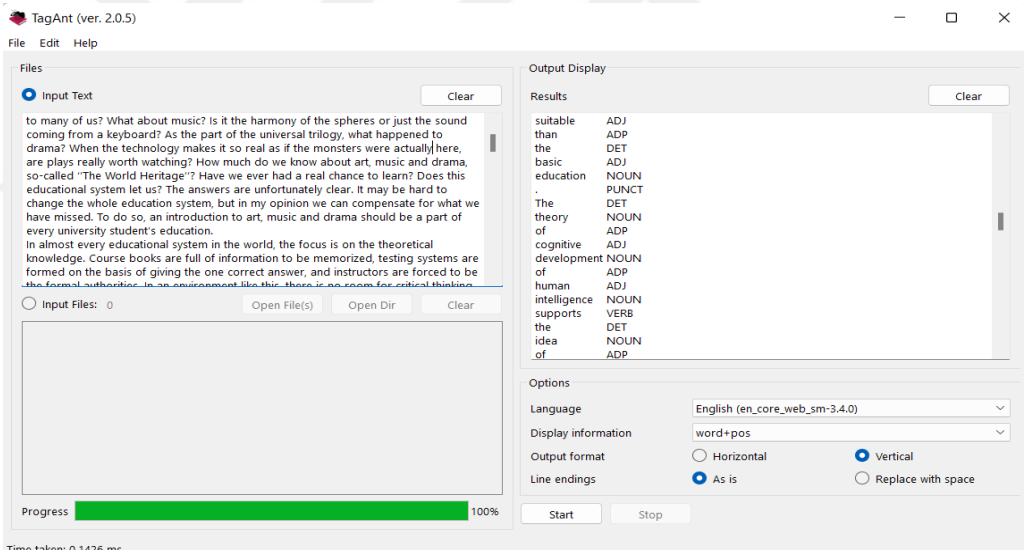
verbs are represented by the small letter "s.". Examples are: *she looks fine; the flowers smell nice*; etc.

3.3.2 Identification of the grammatical collocation patterns

The initial phase involved using *TagAnt* (version 2.0.5), a freeware tagging application created by Laurence Anthony (Anthony, 2022), to identify the Part-Of-Speech (POS) within the learner and native corpus data. This software application takes either an input text or an input list of text files (UTF-8 encoded) and assigns Parts-Of-Speech (POS) tags to them. Initially, learner corpus files were modified to text files and these files were uploaded to the “Input Text” section in the *TagAnt* tool as illustrated in Figure 3.1. The same process was carried out for the native corpus.

Figure 3.1

TagAnt Program

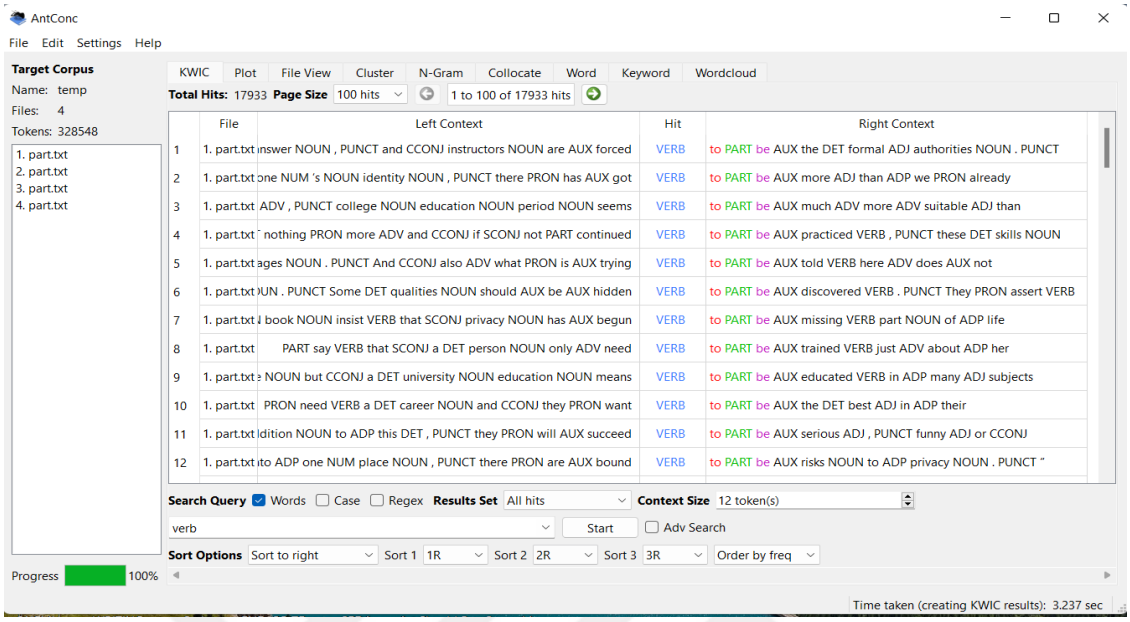


Subsequently, the tagged data was imported into *AntConc* (Version 4.2.4), a corpus analysis toolkit developed by Laurence Anthony (Anthony, 2023). *AntConc* is purpose-built for concordance and text analysis, providing a range of functions such as wordlist creation, keyword identification, concordance analysis, collocate identification, cluster analysis, and N-Gram analysis.

To identify grammatical collocation patterns, the KWIC (Keyword In Context) tab in *AntConc* was used to locate all instances of these patterns as illustrated in Figure 3.2. The results for each pattern were then saved as individual text files.

Figure 3.2

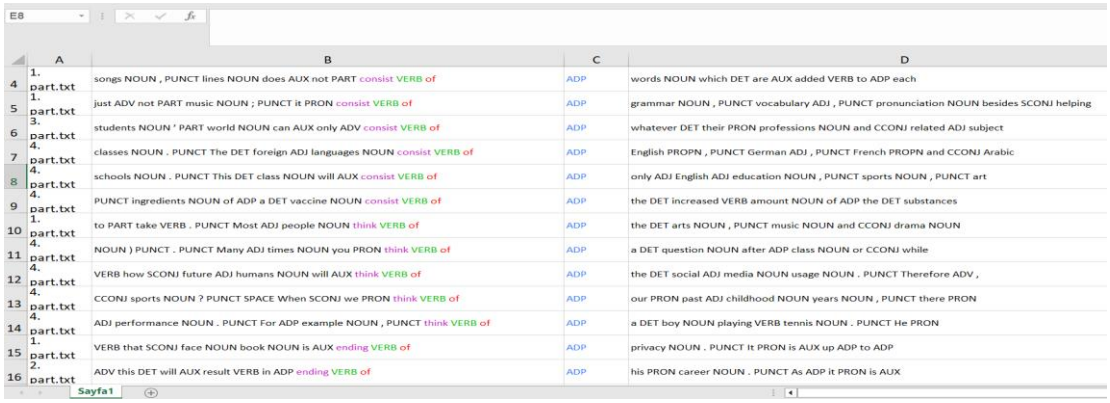
AntConc Program



In the subsequent stage, the text files were transferred into Excel spreadsheets for more detailed analysis (see Figure 3.3). During this review process, occurrences that did not meet the predefined criteria were excluded as irrelevant cases. Ultimately, a final list of grammatical collocation combinations was derived through this iterative process. The same analytical procedure was systematically implemented for both the learner and native corpora to ensure a thorough comparison and evaluation.

Figure 3.3

Text files on an Excel sheet



3.3.3 Data analysis

During the data analysis process, the conclusive set of instances was compiled with reference to the specific research questions guiding the study. The analyses conducted to address each research question of the study drive the core inquiries of the study regarding the types, frequency, and error sources of grammatical collocations among Turkish ELT students and native speakers are as follows:

1. What are the types and frequencies of grammatical collocations appropriately produced by the Turkish ELT students and native speakers?
2. What is the frequency of the grammatical collocational errors within the learner corpus?
3. What are the sources of the grammatical collocational errors within the learner corpus?

To answer *RQ-1*, each occurrence of grammatical collocations in both corpora was found and by using descriptive statistics, the types of collocations were compared based on the classification proposed by Benson et al. (2010). The purpose of this analysis was to identify and compare the types of grammatical collocations used in learner and native language corpus. Subsequently, all occurrences of grammatical collocations identified in both the learner and native speaker corpora were systematically extracted and subjected to quantitative analysis using descriptive statistics, performed through the SPSS statistical software package. This phase of the analysis aimed to explore the frequency patterns of collocational usage among learners in comparison to native speakers. Specifically, it sought to determine whether learners tended to overuse or underuse particular grammatical collocation structures.

In order to answer *RQ-2* and *RQ-3*, an error analysis was conducted, specifically focusing on grammatical collocations. To guide and structure this analysis, Hong et al.'s (2011) adaptation of Gass and Selinker's (2008) Error Analysis Framework was used. The adopted framework provided a systematic approach for generating and analysing the data. The procedure for conducting this error analysis is detailed in Table 3.3, which outlines the steps followed in the study for error analysis.

Table 3.3

Hong et al.'s (2011) Adaptation of Error Analysis Framework

No	Procedure
1	Data generation
2	Identification of errors
3	Classification of errors
4	Quantification of errors
5	Analysis of sources of errors

To compile the dataset, text files processed through *AntConc* were utilised. Each grammatical collocation was analysed in detail, with attention given to both its semantic and grammatical features. For the identification of errors, all concordance lines were reviewed collaboratively by the researcher and a second evaluator—an experienced English instructor and native Turkish speaker. To verify whether the identified issues constituted collocational errors, questionable instances were cross-referenced with entries in the *Longman Dictionary of Contemporary English* (Pearson Education, 2014) and *Cambridge Dictionary* (Cambridge University Press, n.d.). For further confirmation, additional checks were carried out using the *Oxford Collocations Dictionary* (online version) and the *BBI Dictionary of English Word Combinations* (Benson et al., 2010). Finally, for the final judgment of the errors in the learners' essays, another language expert who is a native speaker of English was asked to analyse the data. Both dictionary entries and native speaker judgments were employed in this study, as the assessment of linguistic units can occasionally yield conflicting opinions among native speakers. This necessitates a two-step verification process, where dictionaries and native judgments are consulted to ensure accuracy and consistency in the evaluation of linguistic elements, as discussed by Juknevičienė (2008) and Nesselhauf (2005). The classification and quantification of errors were based on the grammatical collocation patterns that were proposed by Benson et al.'s (2010) classification.

The sources of collocational errors were analysed and categorised by using a framework adapted from Hong et al. (2011) (see Table 3.4). This framework itself drew inspiration from frameworks initially developed by Richards (1974) and Tarone (1981).

Table 3.4

Framework for classifying the sources of collocational errors (Adapted from Hong et al.'s, 2011, Richards', 1974 and Tarone's, 1981) with examples

Strategies	**Major categorisation of sources of errors**	**Sub-categorisation of sources of errors**	**Examples**
Cognitive strategies	Interlingual Transfer	Transfer from L1 (morphological, syntactic, lexical, semantic)	*told to the students* *that/consist from*
	Intralingual Transfer	a) False concept hypothesised	*It was happened*
		b) Overgeneralisation	*might giving*
		c) Incomplete application of rules	*asked we came*
		d) Ignorance of rule restriction	*interested on*
Communicative strategies	Paraphrase	a) Approximation -semantic affinity -morphological and phonological affinities	*hardness to do* *irrespectful to them*

On the whole, the categorisation of sources of errors on the fundamental concepts of cognitive and communicative strategies came after the major classification of sources of collocational errors, namely interlingual transfer, intralingual transfer as well as paraphrase. The three major categorisations of sources of errors were further classified into seven subcategories of sources of collocational errors. It should be noted that the transfer from L1 (morphological, syntactic, lexical, semantic) was adapted from Tarone's (1981) taxonomy and reclassified into the category of interlingual transfer as the present study assigned interlingual transfer to the category of cognitive strategies and it was considered as negative transfer. Additionally, a subcategory of intralingual transfer, specifically 'incomplete application of rules', was incorporated into this framework, drawing from Richards' (1974) study as suggested by Bıçkı (2012).

The categorisation of the sources of errors in collocations according to this framework is as follows:

1. Cognitive Strategies:

Interlingual Transfer: This category encompassed combinations where morphological,

syntactic, lexical, and semantic elements of the first language (L1) were erroneously transferred into the target language (L2). If a combination exhibited aspects of L1, it fell under this category (e.g. *consist from* instead of *consist of*).

Intralingual Transfer: Combinations falling into this category displayed the transfer of L2 characteristics, including false concept hypotheses, overgeneralisation, incomplete application of rules, and ignorance of rule restrictions. If a combination demonstrated the incorrect application of L2 rules, it was categorised in this section (e.g. *interested on* instead of *interested in*).

2. Communicative Strategies:

Paraphrase: This category included combinations that showed semantic affinity, morphological or phonological similarities. These combinations were used as communicative strategies to convey meaning effectively in the L2. For example, if a combination did not contain a correct paraphrase for a word, it fell under this category (e.g. *irrespectful to them* instead of *disrespectful to them*).

In conclusion, the primary aim of this study is to conduct a comprehensive analysis of the sources of errors identified in a learner corpus with regard to grammatical collocations. By examining the nature and frequency of these errors, the study seeks to shed light on a range of language-related challenges commonly faced by EFL learners. These include the misuse of collocational structures, the overuse of certain patterns that may reflect limited lexical variation, and various grammatical inaccuracies associated with collocational use. Through this analysis, the study not only aims to reveal patterns of learner error but also to provide insights into the underlying linguistic and cognitive factors contributing to these issues.

CHAPTER 4 RESULTS

This chapter presents the results of the analyses with reference to the research questions of the study. The initial section (4.1) outlines the types and frequencies of grammatical collocations generated by Turkish ELT students and native speakers. The second section (4.2) presents the frequency of grammatical collocational errors in the learner corpus. In the final section (4.3), the sources of the grammatical collocational errors within the learner corpus are presented in detail by giving the number of occurrences, examples and suggested corrections for each pattern.

In these sections, error rates are categorised as low, moderate, and high; however, since there is no universally standardised scale for these classifications, the categorisation was determined through consensus between the researcher, an experienced English instructor, and a native English speaker. The researcher, an experienced English instructor, and a native English speaker reached a consensus on the categorisation of error rates as low, moderate, and high in these sections, as there is no universally standardised scale for these classifications. This method is consistent with Milton's (2022) suggestion that the frequency and impact of errors should be assessed within the specific context of their studies or teaching environments. This collaborative categorisation ensures that the classification of errors accurately depicts their significance and relevance within the learning context.

4.1 The types and frequencies of grammatical collocations appropriately produced by Turkish ELT students and native speakers

The first research question sought to examine whether the types of grammatical collocations vary between the native speaker corpus and the learner corpus compiled from Turkish ELT students. Descriptive statistics were employed to present the frequency of grammatical collocations in both corpora, providing an overall picture—particularly when reporting numerical data.

In general, the distribution of patterns in both learner and native corpora indicated that grammatical collocations were more frequently used in the native corpus (*f.* 18998) than they were used in the learner corpus (*f.* 17433). The most frequent pattern in the corpora was G8(F) (*Verb + bare infinitive*), with a percentage of 20,32%, which was followed by

G8(D/d) (*Verb* + *(O)* + *Preposition* + *Object*), with a percentage of 20,27%. The patterns G8(P) (*Verb* + *(Object)* + *Adverbial*), G4 (*Preposition* + *Noun*) and G1 (*Noun* + *Preposition*) were moderately used (9,60%, 6,98%, and 6,29%, respectively). The least frequently utilised pattern in both the learner corpus (0,02%) and the native corpus (0,03%) was G8(K) (*Verb* + *a possessive* +*V-ing*). The patterns G8(J) (*Verb* + *Object* + *V-ing*) (Learner, 0,07%; Native, 0,16%) and G8(M) (*Verb* + *Object* + *to be* + *Complement*) (Learner, 0,07%; Native, 0,09%) were the second least often used in the corpora (See Table 4.1).

Table 4.1

Distribution of Patterns in Learner Corpus and Native Corpus

	Pattern	**LEARNER**		**NATIVE**		**Total**	**Total**
		f.	%	*f.*	%	*f.*	%
G1	Noun + Preposition	1109	6.36	1183	6.23	2292	6.29
G2	Noun + to + Infinitive	461	2.64	985	5.18	1446	3.97
G3	Noun + that-clause	193	1.11	375	1.97	568	1.56
G4	Preposition + Noun	1387	7.96	1155	6.08	2542	6.98
G5	Adjective + Preposition	723	4.15	822	4.33	1545	4.24
G6	Adjective + to + Infinitive	346	1.98	448	2.41	803	2.20
G7	Adjective + that-clause	89	0.51	96	0.51	185	0.51
G8(A)	Verb + D O + to + I O = V + I O + D O	174	1.00	182	0.96	356	0.98
G8(B)	Verb + D O + to + I O	57	0.33	77	0.41	134	0.37
G8(C)	Verb + D O +for + I O = Verb + I O + D O	39	0.22	41	0.22	80	0.22
G8(D/d)	Verb + (O) + Preposition + Object	3551	20.37	3835	20.19	7386	20.27
G8(E)	Verb + to + Infinitive	948	5.44	1211	6.37	2159	5.93
G8(F)	Verb + bare infinitive	3940	22.60	3461	18.22	7401	20.32
G8(G)	Verb + V-ing	101	0.58	194	1.02	295	0.81
G8(H)	Verb + Object + to Infinitive	283	1.62	278	1.46	561	1.54
G8(I)	Verb + Object + Infinitive	125	0.72	129	0.68	254	0.70
G8(J)	Verb + Object + V-ing	13	0.07	30	0.16	43	0.12
G8(K)	Verb + a possessive +V-ing	4	0.02	6	0.03	10	0.03
G8(L)	Verb + that-clause	1126	6.46	860	4.53	1986	5.45
G8(M)	Verb + Object + to be + Complement	12	0.07	17	0.09	29	0.08
G8(N)	Verb + Object + Complement	270	1.55	172	0.91	442	1.21
G8(O)	Verb + Object 1 + Object 2	42	0.24	53	0.28	95	0.26
G8(P)	Verb + (Object) + Adverbial	1622	9.30	1875	9.87	3497	9.60
G8(Q)	Verb + (Object) + wh-clause/ wh-phrase	377	2.16	889	4.68	1266	3.48
G8(R)	It + Verb + Object + to Infinitive/that-clause	28	0.16	67	0.35	95	0.26
G8(S/s)	Verb + Complement (Adj./N)	413	2.37	548	2.89	961	2.64
	TOTAL	**17433**	**100**	**18998**	**100**	**36431**	**100**

The results also indicated that within the learner corpora, the most frequent pattern was G8(F) (*Verb + bare infinitive*) with a percentage of 22,60%. It was followed by the G8(D/d) (*Verb + (O) + Preposition + Object*) pattern (20,37%). Concerning the native corpus, the pattern that occurred most often was G8(D/d) (*Verb + (O) + Preposition + Object*) with a percentage of 20,19%. Next in the sequence was G8(F) (*Verb + bare infinitive*) (18,22%).

Considering the objectives of the study, a Pearson Chi-square test was conducted to determine whether the differences in frequencies between the learner and native corpora were significant. The Chi-square test for independence allows researchers to identify significant variations between the frequencies of the categories of two categorical variables (Pallant, 2007). Due to the difference in the size between the native and learner corpus (Native, 168,593; Learner, 172,509), the corpora were standardised to ensure that the analysis is objective by preventing the artificial inflation or under-representation of frequencies. As Biber et al. 1998) stated, standardisation is a method of adjusting the original frequency counts of texts of varying lengths to facilitate accurate comparison. A straightforward and widely used method in corpus linguistics is to adjust the larger corpus to match the size of the smaller corpus in order to assure comparability between corpora of differing sizes (Biber et al., 1998; McEnery and Hardie, 2012). Therefore, the corpora were standardised to 100,000 words in total. In Table 4.2, next to the raw frequencies, relative frequencies derived from standardised tokens are presented.

Table 4.2

Standardised Frequency Distribution of Patterns Across the Corpora

Pattern	LEARNER		NATIVE	
	raw *f.*	relative *f.*	raw *f.*	relative *f.*
G1	1109	643	1183	702
G2	461	267	985	584
G3	193	112	375	222
G4	1387	804	1129	670
G5	723	419	822	488
G6	346	201	457	271
G7	89	52	96	57
G8(A)	174	101	182	108
G8(B)	57	33	77	46
G8(C)	39	23	41	24
G8(D/d)	3551	2058	3835	2275
G8(E)	948	550	1211	718
G8(F)	3940	2284	3461	2053
G8(G)	101	59	194	115
G8(H)	283	164	278	165
G8(I)	125	72	129	77
G8(J)	13	8	30	18
G8(K)	4	2	6	4
G8(L)	1126	653	860	510
G8(M)	12	7	17	10
G8(N)	270	157	172	102
G8(O)	42	24	53	31
G8(P)	1622	940	1875	1112
G8(Q)	377	219	889	527
G8(R)	28	16	67	40
G8(S/s)	413	239	548	325

Note: Relative frequencies were calculated per 100,000 words

The findings from the Pearson Chi-square test demonstrated a statistically significant disparity, with grammatical collocations occurring at a markedly higher frequency in the native speaker corpus than in the learner corpus. ($X^2(25) = 664.624$, $p < .001$). To determine whether there was a significant difference in the use of each type of grammatical collocations by the learners and native speakers, a further analysis was carried out by conducting a Pearson Chi-square test. The results indicated a statistically significant difference between the learner and native speaker groups in their use of 13 distinct types of grammatical collocations (G2, G3, G6, G8(D), G8(E), G8(F), G8(G), G8(L), G8(N), G8(P), G8(Q), G8(R) and G8(S)) (See Table 4.3).

Table 4.3

Chi-Square Results for Each Pattern

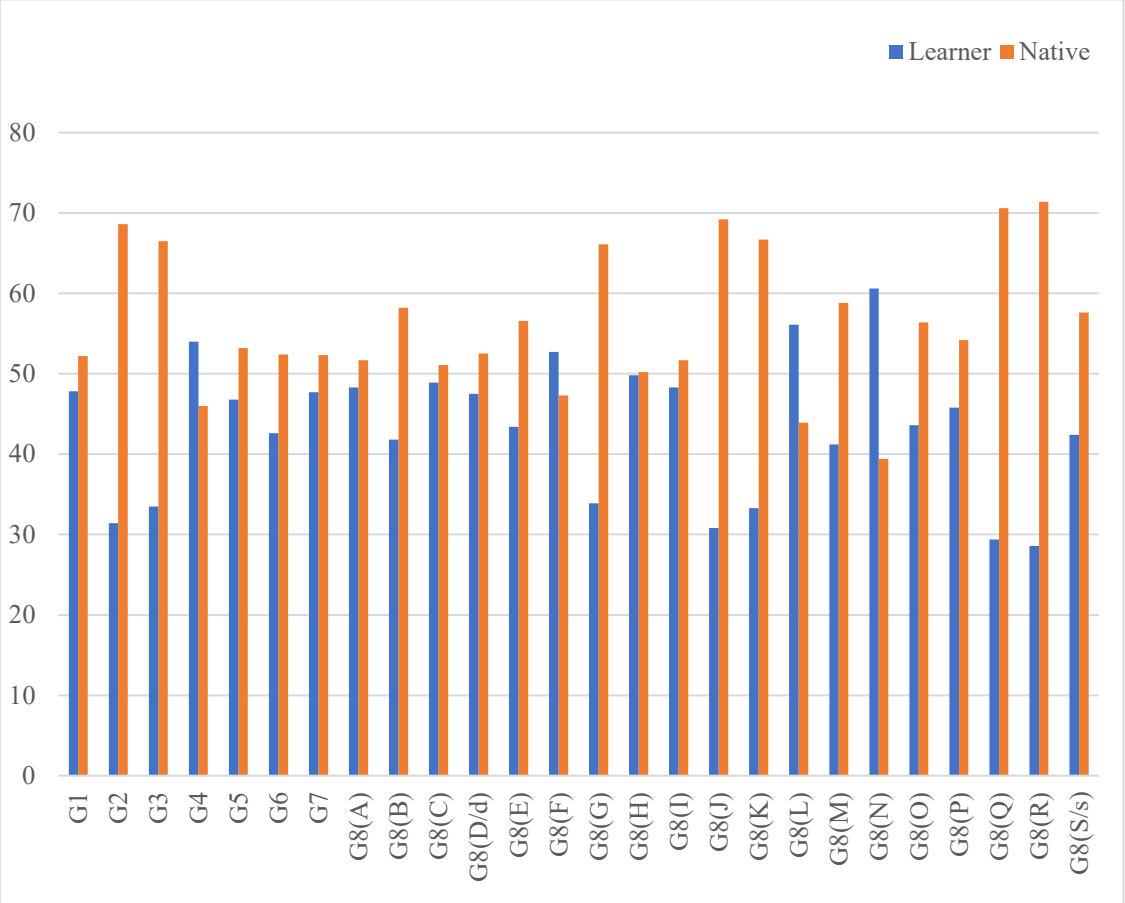
Pattern	LEARNER		NATIVE				
	f.	%	*f.*	%	X^2	df	p
G1	643	47.8	702	52.2	2.59	1	0.18
G2	267	31.4	584	68.6	118	1	< .001
G3	112	33.5	222	66.5	36.2	1	< .001
G4	804	54.0	685	46.0	9.51	1	0.002
G5	419	46.8	488	53.2	6.34	1	0.022
G6	201	42.6	271	52.4	10.4	1	< .001
G7	52	47.7	57	52.3	0.265	1	0.632
G8(A)	101	48.3	108	51.7	0.234	1	0.628
G8(B)	33	41.8	46	58.2	2.99	1	0.144
G8(C)	23	48.9	24	51.1	0.0213	1	0.884
G8(D/d)	2058	47.5	2275	52.5	10.9	1	< .001
G8(E)	550	43.4	718	56.6	22.3	1	< .001
G8(F)	2284	52.7	2053	47.3	12.3	1	< .001
G8(G)	59	33.9	115	66.1	18.0	1	< .001
G8(H)	164	49.8	165	50.2	0.00304	1	0.956
G8(I)	72	48.3	77	51.7	0.168	1	0.682
G8(J)	8	30.8	18	69.2	3.85	1	0.050
G8(K)	2	33.3	4	66.7	0.667	1	0.414
G8(L)	653	56.1	510	43.9	17.6	1	< .001
G8(M)	7	41.2	10	58.8	0.529	1	0.467
G8(N)	157	60.6	102	39.4	11.7	1	< .001
G8(O)	24	43.6	31	56.4	0.891	1	0.345
G8(P)	940	45.8	1112	54.2	14.4	1	< .001
G8(Q)	219	29.4	527	70.6	127	1	< .001
G8(R)	16	28.6	34	71.4	10.03	1	< .001
G8(S/s)	239	42.4	326	57.6	13.1	1	< .001

Note: The percentages indicate the distribution of patterns in the corpora.

As can be seen in Table 4.3 and Figure 4.1, the frequencies of the use of G2 (*Noun + to + infinitive*), G3 (*Noun + that-clause*), G6 (*Adjective + to + Infinitive*), G8(D/d) (*Verb + (O) + Preposition + Object*), G8 (E) (*Verb + to + Infinitive*), G8(G) (*Verb + V-ing*), G8(P) (*Verb + (Object) + Adverbial*), G8(Q) (*Verb + (Object) + wh-clause/ wh-phrase*), G8(R) (*It+ Verb + Object + to Infinitive/that-clause*) and G8(S/s) (*Verb + Complement (Adj. or N)*) patterns were significantly higher in the native corpus than in the learner corpus, whereas the frequencies of G8(F) (*Verb + bare infinitive*), G8(L) (*Verb + that-clause*), G8(N) (*Verb + Object + Complement*) patterns were significantly higher in the learner corpus than in the native corpus.

Figure 4.1

Chi-Square Comparison of Grammatical Collocation Pattern Frequencies in Learner and Native Corpora



Consequently, there was a significant difference between the native corpus and the learner corpus in terms of the use of grammatical collocations. In general, Turkish ELT students utilised grammatical collocations less frequently than native speakers. The results of the current study indicated that prepositional structures (G1 (*noun + preposition*) (Learner, *f.* 1109; Native, *f.* 1183), G5 (*Adjective + Preposition*) (Learner, *f.* 723; Native, *f.* 822) and G8 (D) (*Verb + (Object) + Preposition + Object*) (Learner, *f.* 3551; Native, *f.* 3835) were higher in number in native corpus than in learner corpus. Among these, the G8 (D) pattern demonstrated a statistically significant disparity, occurring with markedly greater frequency in the native corpus. Relatively more complex patterns G8 (Q) (*Verb + (Object) + WH-clause/WH-phrase*) (Learner, *f.* 377; Native, *f.* 889) and G8 (R) (*It + Verb + Object + to-Infinitive/that-clause*) (Learner, *f.* 28; Native, *f.* 67) were also found to be more prevalent in the native corpus, with both patterns occurring at significantly higher frequencies compared to the learner corpus.

Despite notable differences, several similarities were also observed between the learner and native speaker corpora. Both groups demonstrated a frequent use of infinitive and gerund forms, particularly in the patterns G8(E) (*Verb + to + Infinitive*) (Learner: *f.* 948; Native: *f.* 1211) and G8(G) (*Verb + V-ing*) (Learner: *f.* 101; Native: *f.* 194), indicating their central role in English grammar. Additionally, a shared trend was identified in the least frequently used pattern, G8(K) (*Verb + a possessive + V-ing*), which appeared rarely in both corpora (Learner: *f.* 2; Native: *f.* 6).

The findings also revealed that learners exhibited a tendency to overuse certain grammatical collocation patterns. Notably, G8 (F) (*Verb + bare infinitive*) occurred at a significantly higher frequency in the learner corpus. Similarly, G8 (L) (*Verb + that-clause*) was also significantly more frequent among learners compared to native speakers.

4.2 The frequency of grammatical collocational errors within the learner corpus

The second research question focused on identifying the frequency of non-standard or erroneous uses of grammatical collocations within the learner corpus. The results indicated that 681 productions out of 17433 grammatical collocation patterns produced by learners were deviant, which made up 3,91 % of all grammatical collocation combinations. The results also showed that G8(R), *It+ Verb + Object + to Infinitive/that-clause*, had the highest percentage of deviance, 39,29% (See Table 4.4)

Table 4.4

The Overall Distribution of Deviant Uses in the Learner Corpus

	Pattern	Total Occurrences	Deviant Occurrences	%
G1	Noun + Preposition	1109	64	5.77
G2	Noun + to + Infinitive	461	12	2.60
G3	Noun + that-clause	193	15	7.77
G4	Preposition + Noun	1387	45	3.24
G5	Adjective + Preposition	723	30	4.15
G6	Adjective + to + Infinitive	346	7	2.02
G7	Adjective + that-clause	89	10	11.24
G8(A)	Verb + D O + to + I O = V + I O + D O	174	17	9.77
G8(B)	Verb + D O + to + I O	57	7	12.28
G8(C)	Verb + D O +for + I O =Verb + I O + D O	39	4	10.26
G8(D)/(d)	Verb + (Object) + Preposition + Object	3551	184	5.18
G8(E)	Verb + to + Infinitive	948	45	4.75
G8(F)	Verb + bare infinitive	3940	25	0.63
G8(G)	Verb + V-ing	101	12	11.88
G8(H)	Verb + Object + to Infinitive	283	31	10.95
G8(I)	Verb + Object + Infinitive	125	18	14.40
G8(J)	Verb + Object + V-ing	13	3	23.08
G8(K)	Verb + a possessive +V-ing	4	1	25.00
G8(L)	Verb + that-clause	1126	14	1.24
G8(M)	Verb + Object + to be + Complement	12	3	25.00
G8(N)	Verb + Object + Complement	270	22	8.15
G8(O)	Verb + Object 1 + Object 2	42	7	16.67
G8(P)	Verb + (Object) + Adverbial	1622	56	3.45
G8(Q)	Verb + (Object) + wh-clause/ wh-phrase	377	12	3.18
G8(R)	It+ Verb + Object + to Infinitive/that-clause	28	11	39.29
G8(S)/(s)	Verb + Complement (Adj. or N)	413	27	6.54
	TOTAL	**17433**	**681**	**3.91**

Note: The percentages indicate the rate of deviant combinations out of the frequency of each pattern.

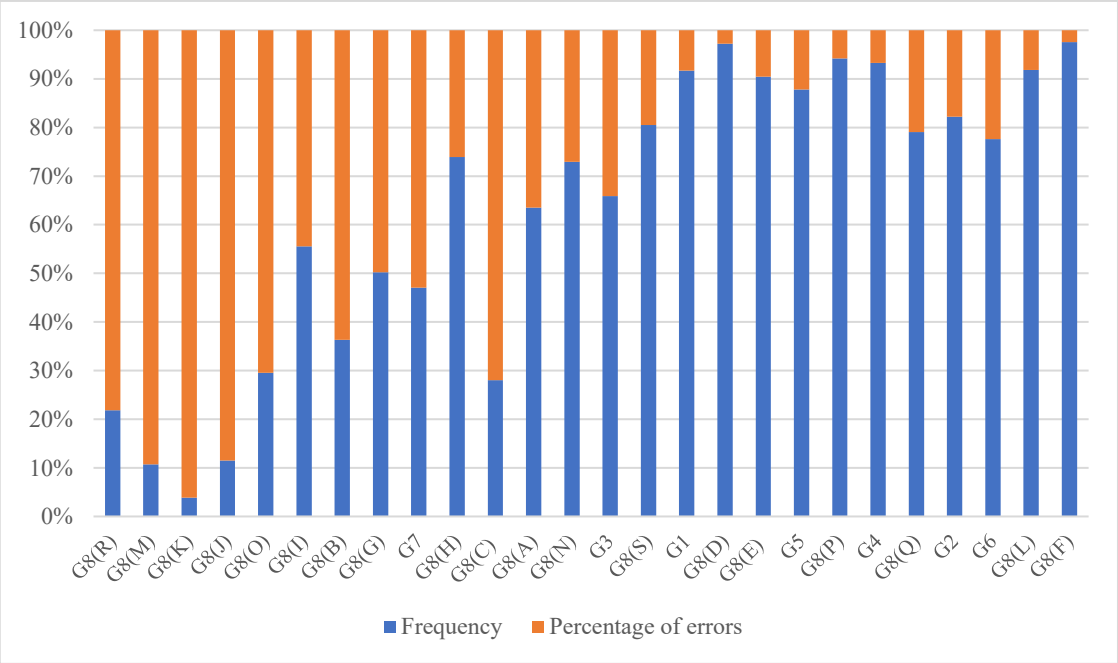
A relatively modest percentage of deviance was observed in patterns G7 (*Adjective + that-clause)* (11,24%), G8(B) (*Verb + D O + to + I O*) (12,28%), G8(C) (*Verb + D O +for + I O =Verb + I O + D O*) (10,26%), G8(G) (*Verb + V-ing)* (11,88%), G8(H) (*Verb + Object + to Infinitive*) (10,95%). The lowest percentage of errors was observed in the use of pattern G8(F), *verb + bare infinitive* (0,7%), which had the highest total frequency, 3940.

It is noteworthy that, certain patterns were not used frequently; however, they had higher percentages of deviance, such as G8(K), *verb + a possessive + v-ing*, and G8(M), *verb + object + to be + complement*, with 25% deviance each, and G8(J), *verb + object + v-ing*, with 23,08% deviance.

Figure 4.2 demonstrates the stacked column chart of deviant uses in the learner corpus ranked. Erroneous uses for each pattern are ranked in order, starting with the pattern which had the highest percentage of deviance (G8(R)) to the lowest (G8(F)).

Figure 4.2

The Stacked Column Chart of Deviant Uses in the Learner Corpus Ranked in Order



It is also notable that there was a relatively lower percentage of deviance in patterns with a high total number of occurrences. To illustrate, the frequency of the use of the G4 (*Preposition + Noun*) pattern was 1387, and the percentage of deviance was 2,52%. Similarly, the G8(L) (*Verb + that-clause*) pattern was relatively common in the corpus with a frequency of 1126, and the percentage of the deviant use of this pattern was 1,24%. A similar case was observed in the use of G8(P) (*Verb + (object) + adverbial).* This pattern had a deviation rate of 3,45% (56 out of 1622).

Consequently, it can be observed that Turkish ELT students encountered difficulties with G8 (K) (Verb + *a possessive + v-ing*), G8(M) (*Verb + object + to be + complement*) and G8(R) *It+ Verb + Object + to Infinitive/that-clause* as these patterns had the highest percentage of errors (25%, 25% and 39,29%, respectively). It is also noteworthy that there were moderate to high deviation rates in patterns with prepositions (e.g., G1 (*Noun + Preposition*) (5,77%), G8 (D/d) (*Verb + (Object) + Preposition + Object*)) (5,18%). Additionally, it was revealed that G8 (F) (*Verb + bare infinitive*) (0,63%) and G8 (L)

(*Verb + that-clause*) (1,24%) showed low deviation, indicating that learners demonstrated proficiency in using *bare infinitives* and *verb + that clauses*.

4.3 The sources of the grammatical collocational errors within the learner corpus

The third research question sought to identify the underlying sources of grammatical collocational errors within the learner corpus. As detailed in the methodology section, the identification of erroneous usages by non-native speakers was carried out through consultation with an experienced Turkish-speaking English instructor, reference to authoritative dictionaries, and validation by a native English speaker. Consequently, a list of deviant combinations for each pattern was compiled. The erroneous uses were examined according to the categorisation of sources of errors on the concepts of *cognitive* (interlingual transfer and intralingual transfer) and *communicative strategies* (paraphrase). The sources of collocational errors were categorised by using a framework adapted from Hong et al. (2011), which was influenced by frameworks originally developed by Richards (1974) and Tarone (1981).

Table 4.5 illustrates the distribution of the total frequency of errors within the major categorisation of sources of errors. The results showed that intralingual errors (63,44%) outnumbered interlingual transfer errors (34,36%). The least frequent source of error was paraphrase with a percentage of 2,20%.

Table 4.5

Distribution of the Total Frequency of Errors within Major Categorisation of Sources of Errors

Major categorisation	**Total *f.***	**%**
Interlingual Transfer	234	34.36
Intralingual Transfer	432	63.44
Paraphrase	15	2.20
	681	**100**

When errors are looked into in terms of the subcategories, it can be observed that with 234 instances, the most common source of error was the *transfer from L1* (34,36%), which is an interlingual transfer error. It was followed by *ignorance of rule restriction* errors within the intralingual transfer category with 143 instances (16,74%). *Incomplete application of rules* was the next most common source of error, with 114 instances spread

across various patterns (9,10%). *False concept hypothesised* error was found in 113 occurrences (16,59%). As for *overgeneralisation,* there were 62 instances of errors (9,25%). As a communicative strategy, within the category of paraphrase, *approximation* errors were relatively rare, with 15 instances across all patterns (2,20%) (See Table 4.6).

Table 4.6

Distribution of the Total Frequency of Errors within Each Category

Strategies	**Major categorisation**	**Sub-categorisation**	**Total *f.***	**%**
	Interlingual Transfer	Transfer from L1 (morphological, syntactic, lexical, semantic)	234	34.36
Cognitive strategies	Intralingual Transfer	False concept hypothesised	113	16.59
		Overgeneralisation	62	9.25
		Incomplete application of rules	114	9.10
		Ignorance of rule restriction	143	16.74
Communicative strategies	Paraphrase	Approximation	15	2.20
		TOTAL	**681**	**100**

The distribution of the *interlingual errors category*, *transfer from L1,* revealed that the highest percentage (40,60%) of this type of error was seen in the pattern G8(D/d) (*Verb + (O) + Preposition + Object*) (*f.* 95). The second highest percentage (11,97%) was observed in patterns G1 (*Noun + Preposition*) (*f.* 28), and G5 (*Adjective + Preposition*) (*f.* 16; 6,84%). Transfer from L1 was rarely observed in patterns G8(R) (*It+ Verb + Object + to Infinitive/that-clause*) (*f.* 2; 0,85%) and G8(H) (*Verb + Object + to Infinitive*) (*f.* 3; 1,28%). There were no instances of *L1 transfer* errors in a total of 12 patterns: G2, G6, G7, G8(C), G8(E), G8(F), G8(G), G8(I), G8(J), G8(K), G8(L), and G8(M). (See Table 4.7).

Table 4.7

Distribution of Interlingual Errors for Each Pattern

Pattern	Transfer from L1	%
G1	28	11.97
G2	0	0.00
G3	7	2.99
G4	14	5.98
G5	16	6.84
G6	0	0.00
G7	0	0.00
G8(A)	14	5.98
G8(B)	6	2.56
G8(C)	0	0.00
G8(D/d)	95	40.60
G8(E)	0	0.00
G8(F)	0	0.00
G8(G)	0	0.00
G8(H)	3	1.28
G8(I)	0	0.00
G8(J)	0	0.00
G8(K)	0	0.00
G8(L)	0	0.00
G8(M)	0	0.00
G8(N)	15	6.41
G8(O)	4	1.71
G8(P)	14	5.98
G8(Q)	8	3.42
G8(R)	2	0.85
G8(S/s)	8	3.42
TOTAL	**234**	**100**

As seen in Table 4.8, the results also revealed that within *intralingual errors*, the most prevalent source of error was *ignorance of rule restriction*, with 143 instances spread across various patterns. The highest frequency was observed in G8 (D/d) (*Verb + (O) + Preposition + Object*) (*f.* 41; 28,67%), G4 (*Preposition + Noun*) (*f.* 29; 20,28%) and G1 (*Noun + Preposition*) (*f.* 20; 13,99%). The lowest frequencies were observed in G8(A) (*Verb + D O + to + I O = V + I O + D O*) and G8(B) (*Verb + D O + to + I O*) (*f.* 1; 0,70%). This type of error was not seen in patterns G2, G3, G6, G7, G8(C), G8(F), G8(G), G8(J), G8(K), G8(M), G8(O), and G8(S/s).

The next frequent source of error was the *incomplete application of rules,* and was prominent in patterns like G8 (D/d) (*Verb + (O) + Preposition + Object*) (f. 37; 32,46%) and G8(E) (*Verb + to + Infinitive*) (*f.* 23; 20,18%). The lowest frequencies regarding this source of error were seen in patterns G8(B) (*Verb + D O + to + I O*) and G8(C) (*Verb + D O +for + I O =Verb + I O + D O*) (*f.* 1; 0,88%). This type of error was not seen in patterns G6, G8(J), G8(K), G8(L), G8(M), G8(N), G8(O), G8(P), and G8(R).

False concept hypothesised error was found in 113 occurrences, with noticeable frequencies in patterns like G8(P) (*Verb + (object) + adverbial)* (*f.* 28; 24,78%) and G1 (*Noun + Preposition*) (*f.* 12; 10,62%). Regarding this pattern, the lowest frequency was observed in pattern G8(K) (*Verb + a possessive +V-ing*) (*f.* 1; 0,88%). There were no instances of this error in patterns G7, G8(A), G8(B) and G8(S).

Within intralingual errors, the least frequent source of error was *overgeneralisation*, with 62 instances of errors, especially in G8(F) (*Verb + bare infinitive*) (*f.* 12; 19,35%) and G8(G) (*f.* 8; 12,90%). Overgeneralisation was not observed in G1, G3, G4, G5, G7, G8(A), G8(B), G8(C), G8(D/d), G8(K), G8(N), G8(O), G8(P) and G8(Q) (See Table 4.8).

Table 4.8

Distribution of Intralingual Errors for Each Pattern

Pattern	*FCH	%	*O	%	*IAR	%	*IRR	%
G1	12	10.62	0	0.00	4	3.51	20	13.99
G2	2	1.77	3	4.84	6	5.26	0	0.00
G3	5	4.42	0	0.00	3	2.63	0	0.00
G4	2	1.77	0	0.00	0	0.00	29	20.28
G5	3	2.65	0	0.00	2	1.75	8	5.59
G6	2	1.77	5	8.06	0	0.00	0	0.00
G7	0	0.00	0	0.00	6	5.26	0	0.00
G8(A)	0	0.00	0	0.00	2	1.75	1	0.70
G8(B)	0	0.00	0	0.00	1	0.88	1	0.70
G8(C)	3	2.65	0	0.00	1	0.88	0	0.00
G8(D/d)	3	2.65	0	0.00	37	32.46	41	28.67
G8(E)	3	2.65	6	9.68	23	20.18	13	9.09
G8(F)	7	6.19	12	19.35	6	5.26	0	0.00
G8(G)	2	1.77	8	12.90	2	1.75	0	0.00
G8(H)	9	7.96	5	8.06	3	2.63	2	1.40
G8(I)	3	2.65	7	11.29	6	5.26	2	1.40
G8(J)	2	1.77	1	1.61	0	0.00	0	0.00
G8(K)	1	0.88	0	0.00	0	0.00	0	0.00
G8(L)	7	6.19	5	8.06	0	0.00	2	1.40
G8(M)	2	1.77	1	1.61	0	0.00	0	0.00
G8(N)	4	3.54	0	0.00	0	0.00	3	2.10
G8(O)	3	2.65	0	0.00	0	0.00	0	0.00
G8(P)	28	24.78	0	0.00	0	0.00	13	9.09
G8(Q)	4	3.54	0	0.00	5	4.39	3	2.10
G8(R)	0	0.00	5	8.06	0	0.00	5	3.50
G8(S/s)	6	5.31	4	6.45	7	6.14	0	0.00
TOTAL	**113**	**100**	**62**	**100**	**114**	**100**	**143**	**100**

** FCH: False Concept Hypothesised, O: Overgeneralisation, IAR: Incomplete application of rules, IRR: Ignorance of rule restriction.*

As for the category of *paraphrase*, *approximation* errors were most frequent in G8(D/d) (*Verb + (O) + Preposition + Object*) with 8 instances (53,33%). There existed only one instance in patterns G1 (*Noun + Preposition*), G2 (*Noun + to + Infinitive*), G5 (*Adjective + Preposition*), G7 (*Adjective + that-clause*) and G8(P) (*Verb + (object) + adverbial),* with a percentage of 6,67. In the remaining patterns G3, G4, G6, G8(A), G8(B), G8(C), G8(E), G8(F), G8(G), G8(H), G8(I), G8(J), G8(K), G8(L), G8(M), G8(N), G8(O), G8(Q), and G8(R) this type of error was not observed (See Table 4.9).

Table 4.9

Distribution of Paraphrase Errors for Each Pattern

Pattern	Approximation	%
G1	1	6.67
G2	1	6.67
G3	0	0.00
G4	0	0.00
G5	1	6.67
G6	0	0.00
G7	1	6.67
G8(A)	0	0.00
G8(B)	0	0.00
G8(C)	0	0.00
G8(D/d)	8	53.33
G8(E)	0	0.00
G8(F)	0	0.00
G8(G)	0	0.00
G8(H)	0	0.00
G8(I)	0	0.00
G8(J)	0	0.00
G8(K)	0	0.00
G8(L)	0	0.00
G8(M)	0	0.00
G8(N)	0	0.00
G8(O)	0	0.00
G8(P)	1	6.67
G8(Q)	0	0.00
G8(R)	0	0.00
G8(S/s)	2	13.33
TOTAL	**15**	**100**

The total number of sources of errors out of 681 indicates that the G8 (D/d) (*Verb + (O) + Preposition + Object*) pattern had the highest total number of errors (184 out of 681), accounting for 27,02% of all errors observed. The pattern G1 *(Noun + Preposition*), with 9,54% of all errors, came next. Other notable error-frequency patterns were G8(P) (*Verb + (Object) + Adverbial*) at 8,22%, G4 (*Preposition + Noun*) at 6,61% and G8(E) (*Verb + to + Infinitive*) at 6,61%, G5 (*Adjective + Preposition*) at 4,41% and G8 (H) (*Verb +*

Object + to Infinitive) at 3,23%. The lowest percentages were observed in G8(J) (Verb + Object + V-ing) (0,44%) and G8(K) (*Verb + a possessive +V-ing*) (0,15%) (See Table 4.10).

The percentages of patterns within each pattern indicated that within transfer of L1, the highest rate was seen in G8 (A) (*Verb + D O + to + I O = V + I O + D O*) (82,35%) and G8 (B) (*Verb + D O + to + I O*) (75%). Regarding false concept hypothesised errors, the highest proportions were found in G8(K) (*Verb + a possessive + V-ing*) at a full 100%, and G8 (C) (*Verb + D O + for + I O = Verb + I O + D O*) at 75%. For overgeneralisation, the most frequent patterns were G6 (*Adjective + to + Infinitive*) at 71,43% and G8 (G) (*Verb + V-ing*) at 66,67%. In the category of incomplete application of rules, high rates were evident in G7 (*Adjective + that-clause*) at 85,71% and G8(E) (*Verb + to + Infinitive*) at 51,11%. Under ignorance of rule restriction, the patterns G4 (*Preposition + Noun*) and G8 (R) (*It + Verb + Object + to Infinitive/that-clause*) showed occurrence rates of 64,44% and 41,67%, respectively. Finally, in terms of approximation, G6 (*Adjective + to + Infinitive*) and G2 (*Noun + to + Infinitive*) displayed the lowest rates at 14,29% and 8.33%, respectively (See Table 4.10).

In the following section (4.3.1), the findings concerning the sources of errors within each collocational pattern are presented. These are organised according to error categories, accompanied by the frequency of occurrence, illustrative examples, and recommended corrections.

Table 4.10

Distribution of Sources of Errors for Each Pattern

	*TL		*FCH		*O		*IAR		*IRR		*A		TOTAL	
Pattern	***f.***	**%**	***f.***	**%**	***f.***	**%**	***f.***	**%**	***f.***	**%**	***f.***	**%**	***f.***	**%**
G1	28	43,08	12	18,46	0	0,00	4	6,15	20	30,77	1	1,54	65	9,54
G2	0	0,00	2	16,67	3	25,00	6	50,00	0	0,00	1	8,33	12	1,76
G3	7	46,67	5	33,33	0	0,00	3	20,00	0	0,00	0	0,00	15	2,20
G4	14	31,11	2	4,44	0	0,00	0	0,00	29	64,44	0	0,00	45	6,61
G5	16	53,33	3	10,00	0	0,00	2	6,67	8	26,67	1	3,33	30	4,41
G6	0	0,00	2	28,57	5	71,43	0	0,00	0	0,00	0	0,00	7	1,03
G7	0	0,00	0	0,00	0	0,00	6	85,71	0	0,00	1	14,29	7	1,03
G8(A)	14	82,35	0	0,00	0	0,00	2	11,76	1	5,88	0	0,00	17	2,50
G8(B)	6	75,00	0	0,00	0	0,00	1	12,50	1	12,50	0	0,00	8	1,17
G8(C)	0	0,00	3	75,00	0	0,00	1	25,00	0	0,00	0	0,00	4	0,59
G8(D/d)	95	51,63	3	1,63	0	0,00	37	20,11	41	22,28	8	4,35	184	27,02
G8(E)	0	0,00	3	6,67	6	13,33	23	51,11	13	28,89	0	0,00	45	6,61
G8(F)	0	0,00	7	28,00	12	48,00	6	24,00	0	0,00	0	0,00	25	3,67
G8(G)	0	0,00	2	16,67	8	66,67	2	16,67	0	0,00	0	0,00	12	1,76
G8(H)	3	13,64	9	40,91	5	22,73	3	13,64	2	9,09	0	0,00	22	3,23
G8(I)	0	0,00	3	16,67	7	38,89	6	33,33	2	11,11	0	0,00	18	2,64
G8(J)	0	0,00	2	66,67	1	33,33	0	0,00	0	0,00	0	0,00	3	0,44
G8(K)	0	0,00	1	100,00	0	0,00	0	0,00	0	0,00	0	0,00	1	0,15
G8(L)	0	0,00	7	50,00	5	35,71	0	0,00	2	14,29	0	0,00	14	2,06
G8(M)	0	0,00	2	66,67	1	33,33	0	0,00	0	0,00	0	0,00	3	0,44
G8(N)	15	68,18	4	18,18	0	0,00	0	0,00	3	13,64	0	0,00	22	3,23
G8(O)	4	57,14	3	42,86	0	0,00	0	0,00	0	0,00	0	0,00	7	1,03
G8(P)	14	25,00	28	50,00	0	0,00	0	0,00	13	23,21	1	1,79	56	8,22
G8(Q)	8	40,00	4	20,00	0	0,00	5	25,00	3	15,00	0	0,00	20	2,94
G8(R)	2	16,67	0	0,00	5	41,67	0	0,00	5	41,67	0	0,00	12	1,76
G8(S/s)	8	29,63	6	22,22	4	14,81	7	25,93	0	0,00	2	7,41	27	3,96
TOTAL	**234**	**100**	**113**	**100**	**62**	**100**	**114**	**100**	**143**	**100**	**15**	**100**	**681**	**100**

Note: The percentages indicate the rate of sources of errors within each pattern.

**TL: Transfer from L1, FCH: False concept hypothesised, O: Overgeneralisation, IAR: Incomplete application of rules, IRR: Ignorance of rule restriction, A: Approximation.*

4.3.1 Detailed categorisation of errors in patterns with samples

The categorisation of error sources within each collocational pattern, including frequency of occurrence, examples of deviant collocations, and recommended corrections, was conducted with the support of an experienced English instructor whose first language is Turkish, a native English speaker, and through consultation of specialised collocation dictionaries. The findings are demonstrated in tables 4.11-4.36.

For pattern G1 (n*oun* + *preposition*), the most frequent type of error was transfer from L1 (*f.* 28). It was followed by ignorance of rule restriction (*f.* 20). False concept hypothesised (*f.* 12) and incomplete application of rules (*f.* 4). There was only one case of approximation error. Overgeneralisation error was not observed (See Table 4.11).

Table 4.11

The Detailed Categorisation of the Errors in Pattern G1 (noun + preposition)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	28	*They will have difficulties at telephone services…*	*They will have difficulties with telephone services…*
False concept hypothesised	12	*Students can concentration on problems more.*	*Students can concentrate on problems more.*
Overgeneralisation	0		
Incomplete application of rules	4	*This can result from deficiency the personality…*	*This can result from deficiency in the personality…*
Ignorance of rule Restriction	20	*They are deterrent factors with crime.*	*They are deterrent factors for crime.*
Approximation	1	*…feelings of inferiority and insatisfaction with her life.*	*… feelings of inferiority and dissatisfaction with her life.*

In Table 4.12, it is presented that within pattern G2 (*noun* + *to* + *infinitive*), the predominant form of error was the incomplete application of rules (*f. 6*). It was preceded by overgeneralisation (*f. 3*), false concept hypothesised (*f.* 2) and approximation (*f.* 1). No instances of transfer from L1, or ignorance of restriction errors were detected.

Table 4.12

The Detailed Categorisation of the Errors in Pattern G2 (noun + to + infinitive)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	0		
False concept hypothesised	2	*They have the enable to think creatively.*	*They have the ability to think creatively.*
Overgeneralisation	3	*…the decision to asking for mental health support…*	*…the decision to ask for mental health support…*
Incomplete application of rules	6	*Efforts help struggling students may include…*	*Efforts to help struggling students may include…*
Ignorance of rule Restriction	0		
Approximation	1	*…due to their wanting to know more about this subject*	*…due to their willingness to know more about this subject*

The primary source of error for pattern G3 (*noun + that clause*) was transfer from L1 (*f.* 7) and it was followed by false concept hypothesised (*f.* 5) and incomplete application of rules (*f.* 3). There were no instances of overgeneralisation, ignorance of rule restriction, or approximation errors (See Table 4.13).

Table 4.13

The Detailed Categorisation of the Errors in Pattern G3 (noun + that clause)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	7	*The reason that because they leave school early is unknown.*	*The reason why they leave school early is unknown.*
False concept hypothesised	5	*The possibility that if online education will become permanent…*	*The possibility that online education will become permanent…*
Overgeneralisation	0		
Incomplete application of rules	3	*The idea that students should study regularly.*	*The idea that students should study regularly is…*
Ignorance of rule Restriction	0		
Approximation	0		

As shown in Table 4.14, ignorance of rule restriction was the most common source of error (*f.* 29) in pattern G4 (*preposition + noun*). It was followed by transfer from L1 (*f.* 14) and 2 cases of false concept hypothesised. There were not any instances of errors in overgeneralisation and incomplete application of rules categories.

Table 4.14

The Detailed Categorisation of the Errors in Pattern G4 (preposition + noun)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	14	*With this way I think the problem can solved.*	*In this way I think the problem can solved.*
False concept hypothesised	2	*The population continues to increase despite of the fact that many young people are moving abroad…*	*The population continues to increase despite the fact that many young people are moving abroad…*
Overgeneralisation	0		
Incomplete application of rules	0		
Ignorance of rule Restriction	29	*Many thought the new policy would decrease productivity; in the contrary, it led to significant improvements.*	*Many thought the new policy would decrease productivity; on the contrary, it led to significant improvements.*
Approximation	0		

Further analysis revealed that in pattern G5 (*adjective + preposition*), except for overgeneralisation, all types of errors were observed. The primary cause of error was transfer from L1 (*f.* 16). It was followed by ignorance of rule restriction (*f.* 8), 3 cases of false concept hypothesised, 2 cases of incomplete application of rules, and only one instance of approximation (See Table 4.15).

Table 4.15

The Detailed Categorisation of the Errors in Pattern G5 (adjective + preposition)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	16	*The additional information was irrelevant with solving the current issue.*	*The information was irrelevant to solving the current issue.*
False concept hypothesised	3	*Many teenagers today are addictive to the internet, spending hours…*	*Many teenagers today are addicted to the internet, spending hours…*
Overgeneralisation	0		
Incomplete application of rules	2	*Being aware the media's influence is essential for developing…*	*Being aware of the media's influence is essential for developing…*
Ignorance of rule restriction	8	*They may feel jealous with their posts and…*	*They may feel jealous of their posts and…*
Approximation	1	*The new policy became uneffective in reducing crime rates.*	*The new policy became ineffective in reducing crime rates.*

In general, errors were infrequent for pattern G6 (*adjective + to + infinitive*). The major reason for inaccuracy was overgeneralisation (*f.* 5). Another source of inaccuracy was the false concept hypothesis (*f.* 2). In the remaining categories (Transfer from L1, incomplete application of rules, ignorance of rule restriction and approximation), errors were not detected (See Table 4.16).

Table 4.16

The Detailed Categorisation of the Errors in Pattern G6 (adjective + to + infinitive)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	0		
False concept hypothesised	2	*It is importance to be prepared for challenges in life.*	*It is important to be prepared for challenges in life.*
Overgeneralisation	5	*New vocabulary is often easier to learning when used in real-life contexts.*	*New vocabulary is often easier to learn when used in real-life contexts.*
Incomplete application of rules	0		
Ignorance of rule Restriction	0		
Approximation	0		

The detailed classification of the errors in pattern G7 (*adjective + that clause*) suggested that errors were not frequent in general. The primary cause of error was incomplete application of rules (*f.* 6). There was only one case of approximation (*f.* 1). Other types of errors (Transfer from L1, false concept hypothesised, overgeneralisation and ignorance of rule restriction) were not detected (See Table 4.17).

Table 4.17

The Detailed Categorisation of the Errors in Pattern G7 (adjective + that clause)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	0		
False concept hypothesised	0		
Overgeneralisation	0		
Incomplete application of rules	6	*It is certain that will cause problems.*	*It is certain that it will cause problems.*
Ignorance of rule Restriction	0		
Approximation	1	*This situation seems undisdutable.*	*This situation seems indisputable.*

Based on pattern G8(A) (*Verb + D O + to + I O = V + IO + D O*), the most frequent source of error was transfer from L1 (*f.* 14). It was followed by incomplete application of rules (*f.* 2). Only one instance of ignorance of rule restriction error was observed. No other sources of errors (False concept hypothesised, overgeneralisation and approximation) were detected (See Table 4.18).

Table 4.18

The Detailed Categorisation of the Errors in Pattern G8(A) (Verb + D O + to + I O = V + IO + D O)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	14	*Teachers prefer to give to students feedback*	*Teachers prefer to give students feedback…/ Teachers prefer to give feedback to the students*
False concept hypothesised	0		
Overgeneralisation	0		
Incomplete application of rules	2	*They sent these photos each other to make a collaborative…*	*They sent these photos to each other to make a collaborative…*
Ignorance of rule Restriction	1	*Local businesses offer a lot of savings of students on entertainment…*	*Local businesses offer a lot of savings to students on entertainment…*
Approximation	0		

As presented in Table 4.19, for pattern G8(B) (*Verb + D O + to + I O*), transfer from L1 was the most common source of error (*f.* 6). It was followed one case of incomplete application of rules error and one case of ignorance of rule restriction error. False concept hypothesised, overgeneralisation and approximation errors were not observed.

Table 4.19

The Detailed Categorisation of the Errors in Pattern G8(B) (Verb + D O + to + I O)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	6	*It is essential for students give attention to the lesson…*	*It is essential for students to pay attention to the lesson…*
False concept hypothesised	0		
Overgeneralisation	0		
Incomplete application of rules	1	*Then, teachers explain the project to.*	*Then, teachers explain the project to the students.*
Ignorance of rule Restriction	1	*The use of digital media brings different dimensions at communication*	*The use of digital media brings different dimensions to communication*
Approximation	0		

Table 4.20 shows that learners generated a total of 3 false concept hypothesis errors and 1 error of incomplete application of rules for pattern G8(C) (*Verb + D O +for + I O =Verb + I O + D O*). There were no instances of transfer from L1, overgeneralisation, ignorance of rule restriction, or approximation errors.

Table 4.20

The Detailed Categorisation of the Errors in Pattern G8(C) (Verb + D O +for + I O = Verb + I O + D O)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	0		
False concept hypothesised	3	*Some believe society should show more respect on famous people.*	*Some believe society should show more respect to famous people.*
Overgeneralisation	0		
Incomplete application of rules	1	*They prepare a lesson plan their students.*	*They prepare a lesson plan for their students. /They prepare their students a lesson plan.*
Ignorance of rule Restriction	0		
Approximation	0		

Table 4.21 indicates that learners produced 184 errors for pattern G8(D)/(d) (*Verb + (Object) + Preposition + Object*). The bulk of these were instances of transfer from L1 (*f.* 95). Another common category was ignorance of rule restriction, with 41 instances. It was followed by incomplete application of rules errors (*f.* 37), approximation (*f.* 8), false concept hypothesised (*f.* 3.). There were no instances of overgeneralisation.

Table 4.21

The Detailed Categorisation of the Errors in Pattern G8(D)/(d) (Verb + (Object) + Preposition + Object)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	95	*They like to listen music while studying.*	*They like to listen to music while studying.*
False concept hypothesised	3	*Students are encouraged to take responsible for their learning.*	*Students are encouraged to take responsibility for their learning.*
Overgeneralisation	0		
Incomplete application of rules	37	*Before making a big decision, it's essential to think about.*	*Before making a big decision, it's essential to think about its results.*
Ignorance of rule Restriction	41	*The residents began to complain of this situation…*	*The residents began to complain about this situation…*
Approximation	8	*We should stop ourselves from praising actions that go against our values.*	*We should prevent ourselves from praising actions that go against our values.*

The most common type of error in pattern G8(E) (*Verb + to + Infinitive*) was incomplete application of rules (*f.* 23). It was then preceded by ignorance of restriction (*f.* 13), overgeneralisation (*f.* 6) and false concept hypothesised (*f.* 3). Transfer from L1 or approximation were not detected (See Table 4.22).

Table 4.22

The Detailed Categorisation of the Errors in Pattern G8(E) (Verb + to + Infinitive)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	0		
False concept hypothesised	3	*Their mentor suggested to work independently to… develop their problem-solving skills.*	*Their mentor wanted them to work independently to develop their problem-solving skills.*
Overgeneralisation	6	*Experts suggest to study on a particular subject.*	*Experts suggest studying a particular subject.*
Incomplete application of rules	23	*They don't need to anxious about the outcome, as everything is under control.*	*They don't need to be anxious about the outcome, as everything is under control.*
Ignorance of rule Restriction	13	*The criminal pretended for being someone else to escape…*	*The criminal pretended to be someone else to escape…*
Approximation	0		

As shown in Table 4.23, for pattern G8(F) *(Verb + bare infinitive),* the most frequent type of error was overgeneralisation (*f.* 12). It was followed by false concept hypothesis (*f.* 7) and incomplete application of rules (*f.* 6). The unobserved categories were transfer from L1, ignorance of rule restriction and approximation.

Table 4.23

The Detailed Categorisation of the Errors in Pattern G8(F) (Verb + bare infinitive)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	0		
False concept hypothesised	7	*Not getting enough sleep may be cause difficulties with concentration and memory.*	*Not getting enough sleep may cause difficulties with concentration and memory.*
Overgeneralisation	12	*Building good habits will helps them succeed…*	*Building good habits will help them succeed…*
Incomplete application of rules	6	*Some people hesitate share their ideas in meetings, fearing criticism.*	*Some people hesitate to share their ideas in meetings, fearing criticism.*
Ignorance of rule Restriction	0		
Approximation	0		

The detailed categorisation of the errors in pattern G8(G) *(Verb + V-ing)* shows that overgeneralisation was the most frequent source of error (*f.* 8). There were 2 cases of false concept hypothesis and incomplete application of rules error, and no instances of other types (Transfer from L1, ignorance of rule restriction and approximation) (See Table 4.24).

Table 4.24

The Detailed Categorisation of the Errors in Pattern G8(G) (Verb + V-ing)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	0		
False concept hypothesised	2	*Teachers often need to guiding students about complex topics.*	*Teachers often need to guide students about complex topics.*
Overgeneralisation	8	*After careful consideration, he decides sending additional resources…*	*After careful consideration, he decides to send additional resources…*
Incomplete application of rules	2	*They insist using traditional methods in…*	*They insist on using traditional methods in…*
Ignorance of rule Restriction	0		
Approximation	0		

As presented in Table 4.25, for pattern G8 (H) (*Verb + Object + to Infinitive*), except for approximation error, all sources of errors were observed. The most frequent type of error was the false concept hypothesis, with an occurrence of 9. It was followed by overgeneralisation (*f.* 5), incomplete application of rules (*f.* 3) and transfer from L1 (*f.* 3). The least frequent type of error was ignorance of rule restriction (*f.* 2).

Table 4.25

The Detailed Categorisation of the Errors in Pattern G8(H) (Verb + Object + to Infinitive)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	3	*The new program provided students to take advanced courses.*	*The new program enabled students to take advanced courses.*
False concept hypothesised	9	*The new law permits agencies to listening to private conversations.*	*The new law permits agencies to listen to private conversations.*
Overgeneralisation	5	*Strict regulations can sometimes prevent people to live in certain areas.*	*Strict regulations can sometimes prevent people from living in certain areas.*
Incomplete application of rules	3	*They want their children have opportunities.*	*They want their children to have opportunities.*
Ignorance of rule Restriction	2	*Social media helps people for connect with friends and family*	*Social media helps the people to connect with friends and family…*
Approximation	0		

For pattern G8(I) (*Verb + Object + Infinitive*), a total of 17 errors were produced by learners. There were 7 cases of overgeneralisation. Next in sequence were 6 cases of incomplete application of rules, 3 cases of false concept hypothesised, and ignorance of rule restriction with 2 cases. There were no cases of transfer from L1 or approximation (See Table 4.26).

Table 4.26

The Detailed Categorisation of the Errors in Pattern G8(I) (Verb + Object + Infinitive)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	0		
False concept hypothesised	3	*Helping others makes feel them better and brings…*	*Helping others makes them feel better and brings…*
Overgeneralisation	7	*They planned a hands-on activity to let them to discover new concepts on their own.*	*They planned a hands-on activity to let them discover new concepts on their own.*
Incomplete application of rules	6	*Various methods let improve their talents in…*	*Various methods let them improve their talents in…*
Ignorance of rule Restriction	2	*Regular cleaning can cause germs and toxins for fade away…*	*Regular cleaning can cause germs and toxins to fade away…*
Approximation	0		

As for pattern G8(J) *(Verb + Object + V-ing),* errors were not frequent. There were two cases of false concept hypothesised error and only one case of overgeneralisation. The unobserved categories were transfer from L1, incomplete application of rules, ignorance of rule restriction and approximation (See Table 4.27).

Table 4.27

The Detailed Categorisation of the Errors in Pattern G8(J) (Verb + Object + V-ing)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	0		
False concept hypothesised	2	*By dematerializing them thinking, individuals focus more on ideas, experiences, and values.*	*By dematerializing their thinking, individuals focus more on ideas, experiences, and values.*
Overgeneralisation	1	*Professionals who frequently keep their clients wait risk their reputation and…*	*Professionals who frequently keep their clients waiting risk their reputation and…*
Incomplete application of rules	0		
Ignorance of rule Restriction	0		
Approximation	0		

In Table 4.28, it can be seen that pattern G8(K) (Verb + possessive + V-ing) contained only one error, and its source was false concept hypothesised. The other sources of errors (Transfer from L1, overgeneralisation, incomplete application of rules, ignorance of rule restriction and approximation) were not detected.

Table 4.28

The Detailed Categorisation of the Errors in Pattern G8(K) (Verb + possessive + V- ing)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	0		
False concept hypothesised	1	*Teachers dislike their students' arrive late to the class.*	*Teachers dislike their students' arriving late to the class.*
Overgeneralisation	0		
Incomplete application of rules	0		
Ignorance of rule Restriction	0		
Approximation	0		

For pattern G8(L) (*Verb + that-clause*), false concept hypothesis was the most frequent source, with 7 occurrences. It was followed by overgeneralisation error (*f.* 5). There were 2 cases of ignorance of rule restriction error. The types of errors that were not observed were transfer from L1, incomplete application of rules, ignorance of rule restriction and approximation (See Table 4.29).

Table 4.29

The Detailed Categorisation of the Errors in Pattern G8(L) (Verb + that-clause)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	0		
False concept hypothesised	7	*A lack of practice often leads that students do not to acquire the necessary skills…*	*A lack of practice often leads students not to acquire the necessary skills…*
Overgeneralisation	5	*They are discussing that whether there should be new regulations for environmental protection.*	*They are discussing whether there should be new regulations for environmental protection.*
Incomplete application of rules	0		
Ignorance of rule Restriction	2	*They assert which there is no need for additional funding.*	*They assert that there is no need for additional funding.*
Approximation	0		

In pattern G8(M) *(Verb + Object + to be + Complement)*, errors were not frequent, with a total of 3 cases. 2 of them were false concept hypothesised errors and only one was an overgeneralisation error. There were no cases of Transfer from L1, incomplete application of rules, ignorance of rule restriction, or approximation errors (See Table 4.30).

Table 4.30

The Detailed Categorisation of the Errors in Pattern G8(M) (Verb + Object + to be + Complement)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	0		
False concept hypothesised	2	*Parents want their children to be succeed at school.*	*Parents want their children to be successful at school.*
Overgeneralisation	1	*Experts suggest be done with caution for safety.*	*Experts suggest it to be done with caution for safety.*
Incomplete application of rules	0		
Ignorance of rule Restriction	0		
Approximation	0		

As presented in Table 4.31, the detailed categorisation of the errors in pattern G8(N) *(Verb + Object + Complement)* indicates that transfer from L1 was by far the most frequent source of error (*f.* 15). The other cases were sourced as false concept hypothesised (*f.* 4) and ignorance of rule restriction (*f.* 3). Overgeneralisation, incomplete application of rules, and approximation errors were not detected.

Table 4.31

The Detailed Categorisation of the Errors in Pattern G8(N) (Verb + Object + Complement)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	15	*Teachers make easier understand by using different methods.*	*Teachers make it easier to understand by using different methods.*
False concept hypothesised	4	*Society considers important them because their contributions positively impact the community.*	*Society considers them important because their contributions positively impact the community.*
Overgeneralisation	0		
Incomplete application of rules	0		
Ignorance of rule Restriction	3	*Making courses for compulsory ensures that all students gain essential knowledge and skills.*	*Making courses compulsory ensures that all students gain essential knowledge and skills.*
Approximation	0		

For G8(O) *(Verb + Object 1 + Object 2)*, there were 4 cases of transfer from L1 and 3 cases of false concept hypothesised error. The unobserved categories were overgeneralisation, incomplete application of rules, ignorance of rule restriction, and approximation (See Table 4.32).

Table 4.32

The Detailed Categorisation of the Errors in Pattern G8(O) (Verb + Object 1+ Object 2)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	4	*In traditional classroom settings, education requires students attendance.*	*In traditional classroom settings, education requires students' attendance.*
False concept hypothesised	3	*Experts give to them the necessary information.*	*Experts give them the necessary information.*
Overgeneralisation	0	*They provide parents information in certain situations.*	*They provide parents with information in certain situations.*
Incomplete application of rules	0		
Ignorance of rule Restriction	0		
Approximation	0		

The detailed categorisation of the errors in pattern G8(P) *(Verb + (Object) + Adverbial)* shows that the most frequent error was false concept hypothesised, with 28 occurrences. It was followed by transfer from L1, with a total of 14 cases, and 13 cases of ignorance of rule restriction error. Only one instance of an approximation error was detected. Overgeneralisation and incomplete application of rules errors were not observed (See Table 4.33).

Table 4.33

The detailed categorisation of the errors in pattern G8(P) (Verb + (Object) + Adverbial)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	14	*Children can remember beautifully the moments when their parents spent quality time with them.*	*Children can remember fondly the moments when their parents spent quality time with them.*
False concept hypothesised	28	*The advantages of a new system cannot be achieved complete without proper planning.*	*The advantages of a new system cannot be achieved completely without proper planning.*
Overgeneralisation	0		
Incomplete application of rules	0		
Ignorance of rule Restriction	13	*If environmental awareness continues to grow, the world will more responsibly act to protect future generations.*	*If environmental awareness continues to grow, the world will act more responsibly to protect future generations.*
Approximation	1	*They are expected to act more thinkfully than others.*	*They are expected to act more thoughtfully than others.*

As presented in Table 4.34, for G8(Q) *(Verb + (Object) + wh-clause/ wh-phrase)*, the most frequent source of error was transfer from L1 with a total of 8 errors. It was followed by 5 instances of incomplete application of rules, 4 instances of false concept hypothesised, and 3 instances of ignorance of rule restriction errors. The sources of errors that were not detected were overgeneralisation and approximation.

Table 4.34

The Detailed Categorisation of the Errors in Pattern G8(Q) (Verb + (Object) + wh- clause/ wh-phrase)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	8	*Students ask them where can they stay.*	*Students ask them where they can stay.*
False concept hypothesised	4	*The evidence doesn't explain why pursuit extensive tools.*	*The evidence doesn't explain why they pursue extensive tools.*
Overgeneralisation	0		
Incomplete application of rules	5	*The software organizes videos by predicting they belong.*	*The software organizes videos by predicting where they belong.*
Ignorance of rule Restriction	3	*Doctors ask them that they are against taking the medication.*	*Doctors ask them whether they are against taking the medication.*
Approximation	0		

For pattern G8(R) (*It+ Verb + Object + to Infinitive/that-clause*), an equivalent number of errors was produced in terms of overgeneralisation and ignorance of rule restriction (*f.* 5). It was followed by 2 occurrences of transfer from L1. No further sources of errors (false concept hypothesised, incomplete application of rules, and approximation) were detected (See Table 4.35).

Table 4.35

The Detailed Categorisation of the Errors in Pattern G8(R) (It+ Verb + Object + to Infinitive/that-clause)

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	2	*The new online platform is beneficial because it provides us to receive training from experts…*	*The new online platform is beneficial because it allows us to receive training from experts…*
False concept hypothesised	0		
Overgeneralisation	5	*Meditation is a powerful practice because it helps us relaxing.*	*Meditation is a powerful practice because it helps us to relax.*
Incomplete application of rules	0		
Ignorance of rule Restriction	5	*Lack of communication often leads to misunderstandings, and it causes to break up.*	*Lack of communication often leads to misunderstandings, and it causes people to break up.*
Approximation	0		

As shown in Table 4.36, transfer from L1 was the most dominant source of error in terms of G8(S)/(s) *(Verb + Complement (Adj. or N))* with an occurrence of 8. A rather less frequent type was incomplete application of rules (*f.* 7). It was followed by 6 cases of false concept hypothesised, 4 cases of overgeneralisation, and 2 cases of approximation. The only unobserved category was ignorance of rule restriction.

Table 4.36

The Detailed Categorisation of the Errors in Pattern G8(S)/(s) (Verb + Complement (Adj. or N))

Sources of errors	Number of occurrences	Examples of deviant collocations	Suggested correction
Transfer from L1	8	*People who work long hours without breaks are more likely to feel stressful.*	*People who work long hours without breaks are more likely to feel stressed.*
False concept hypothesised	6	*If people be more careful about supporting artists, it can help enrich culture.*	*If people are more careful about supporting artists, it can help enrich culture.*
Overgeneralisation	4	*It's important to stay calm when another person feel panicked.*	*It's important to stay calm when another person feels panicked.*
Incomplete application of rules	7	*By joining clubs, one can become acquaint with lots of people.*	*By joining clubs, one can become acquainted with lots of people.*
Ignorance of rule Restriction	0		
Approximation	2	*Children seem inproductive if tasks are not assigned according to their interests.*	*Children seem unproductive if tasks are not assigned according to their interests.*

In conclusion, upon examining the sources of grammatical collocational errors by major categorisation, *intralingual transfer* errors outnumbered *interlingual transfer* errors, and the *paraphrase* error was the least frequent. As for the sub-categorisation of errors, it is evident that the predominant cause of error identified was the transfer from L1, classified as an *interlingual transfer* error. This error was prevalent in patterns such as G8 (D/d) (*Verb + (O) + Preposition + Object*) and G1 (*Noun + Preposition*). As for the *intralingual transfer* category, there was a notable prevalence of ignorance of rule restriction errors, most frequent in patterns G8(D/d), G4 (*Preposition + Noun*) and G1 (*Noun + Preposition*), and G5 (*Adjective + Preposition*). *Incomplete application of rules* was also a prevalent source of errors across many patterns. The greatest frequency was noted in G8 (D/d) and G8(E) (*Verb + to + Infinitive*). Errors in the *false concept*

hypothesised were the next most common in the learner corpus, with significant frequencies observed in patterns such as G8(P) (*Verb + (Object) + Adverbial*) and G1 (*Noun + Preposition*). Regarding *overgeneralisation*, errors were particularly in G8(F) (*Verb + bare infinitive*). Specifically, within paraphrasing, approximation errors were infrequent, occurring most frequently in G8(D/d).

CHAPTER 5 CONCLUSION, DISCUSSION AND SUGGESTIONS

5.1 Conclusion

This study primarily aimed to evaluate the use of grammatical collocations in the argumentative writings of Turkish ELT students whose level of proficiency ranges between the B1 and B2 levels. It examined the natural use of grammatical collocations in a learner corpus using a native speaker corpus. In alignment with the objectives, the study sought to investigate the varieties and frequency of grammatical collocations generated by Turkish ELT students and native speakers, the forms of grammatical collocational errors that exist within the learner corpus and the origins of the errors in the learner corpus, based on the categories of grammatical collocations.

The current study compared the types and frequencies of grammatical collocations used by Turkish learners of English and native speakers of English, aiming to identify and compare the types of collocations used in the learner and native language corpus. The study also aimed to determine which aspect of collocation instruction should be emphasised more in the ELT classroom. In addition to this, a detailed error analysis of grammatical collocation patterns in the learner corpus was presented, identifying types of errors and problematic areas. The study focused on the origins of errors under the categorisation of interlingual transfer, intralingual transfer and communication, aiming to explain issues and challenges associated with learning collocations.

A corpus-based methodology was utilised to analyse the occurrence of grammatical collocations in the natural language use of English learners and native speakers. In order to assess whether learners overuse or underuse these specific combinations, a chi-square test was conducted. The study also employed descriptive analyses, including frequencies and percentages, to clarify the distinctions between the two corpora. Additionally, an interpretative approach was used to examine any instances of incorrect productions in the learner corpus, taking into account predefined categorisations, which is a reliable approach to accurately describe language structure and acquisition, rather than attempting to discover new linguistic features.

As there is a notable scarcity of studies explicitly examining grammatical collocations used by non-native Turkish speakers in Turkish contexts, this study is unique in that it compares the grammatical collocations of non-native speakers based on a corpus of native

speakers' argumentative writings in the Turkish setting. Therefore, the study aimed to fill a gap in the existing literature and offer more precise recommendations for curriculum design, instructional strategies, and material development that will more effectively support learners' overall language proficiency and collocational competence.

A brief overview of the findings related to the research questions is provided below by summarising the key findings:

1. *What are the types and frequencies of grammatical collocations appropriately produced by Turkish ELT students and native speakers?*

The research demonstrated notable disparities in the utilisation of grammatical collocations between the native and learner corpora. Compared to native speakers, Turkish ELT students employed grammatical collocations less frequently overall. In the native corpus, complex structures like G8(Q) (*Verb + (Object) + WH-clause/ WH-phrase*) and G8(R) (*It + Verb + Object + to Infinitive/that-clause*) were far more prevalent. The native corpus also had a higher prevalence of prepositional structures, such as G1 (*Noun + Preposition*), G5 (*Adjective + Preposition*), and several G8 patterns (e.g., G8(A), G8(B), G8(C), and G8(D)). G8(D) (*Verb + (Object) + Preposition + Object*) showed a particularly significant difference. On the other hand, learners frequently overused specific patterns. For instance, G8(F) (*Verb + bare infinitive*) and G8(L) (*Verb + that-clause*) were significantly more prevalent in the learner corpus. Notwithstanding these distinctions, some similarities were observed: both groups regularly employed infinitive and gerund forms, whereas the less prevalent structure G8(K) (*Verb + a possessive + V-ing*) was utilised with equal frequency by both learners and native speakers. These findings underscore the discrepancies in learner collocational usage compared to native standards, as well as instances of overlaps.

Consequently, these findings also indicate that mastering the extensive array of collocational patterns employed by native speakers is frequently a challenge for even the most advanced English learners. They may proficiently employ high-frequency collocations that are frequently taught in language courses; however, they frequently refrain from employing less common, strongly associated collocations that contribute to the idiomatic richness of the language. This reliance on predictable language frequently leads to writing and discourse that lack the naturalness that is indicative of native usage.

Learners frequently encounter difficulties in identifying appropriate word pairings, resulting in unnatural and ungainly combinations (Nesselhauf, 2003; Liu, 2010).

2. *What is the frequency of the grammatical collocational errors within the learner corpus?*

The learner corpus revealed that Turkish ELT students experienced challenges with G8(K) (*Verb + possessive + v-ing*), G8(M) (*Verb + object + to be + complement*), and G8(R) (*It + Verb + Object + to Infinitive/that-clause*), as these structures exhibited the highest error rates. Significantly, there were moderate to high deviation rates in patterns using prepositions (e.g., *G1 (Noun + Preposition*), G8 (D/d) (*Verb + (Object) + Preposition + Object*)). Additionally, the low deviation of several patterns (e.g., G8 (F) (*Verb + bare infinitive*) and G8 (G) (*Verb + V-ing*)) showed that learners were proficient in employing gerund forms and bare infinitives.

The low deviation rate observed for *verb + bare infinitive* and *verb + v-ing* is likely due to this accessibility, which indicates a substantial comprehension of bare infinitives. According to DeKeyser (2005) learners are more likely to acquire structures with explicit, consistent norms, particularly when reinforced through recurrent exposure; therefore, due to their uncomplicated structure, *verb + bare infinitive* and *verb + v-ing* are considered to be highly accessible and simpler for learners to internalise, as they are frequently encountered with modal verbs and causal forms. Nevertheless, learners encountered challenges with more difficult structures, such as G8(K), G8 (M), and G8(R), despite their moderate proficiency in English, as a result of their contextual appropriateness and flexibility (Bourke, 2007). These patterns were frequently overgeneralised or simplified by learners, which led to inaccuracies and misconduct. Additional challenges were presented by grammatical collocations that involve prepositions, as errors occur at a relatively moderate frequency. Prepositions in English are considered to be challenging for even the most advanced learners, as their usage is largely arbitrary and necessitates extensive practice to master (Liao and Fukuya, 2004). In addition, Šebestová (2021) emphasised that the analytical nature of English prepositions and their role in text organisation present additional challenges, particularly for learners with limited exposure to authentic language use or those whose native languages have significantly different prepositional systems. Consequently, these results underscore the discrepancy between learners' capacity to accurately employ simpler structures, such as bare infinitives, and

more intricate grammatical patterns and prepositional collocations. They also emphasise the significance of offering targeted instruction and extensive practice to assist learners in mastering these more difficult aspects of English grammar.

3. *What are the sources of the grammatical collocational errors within the learner corpus?*

The research analysed the sources of grammatical collocational errors within the learner corpus, indicating that *intralingual transfer* errors occurred more frequently than *interlingual transfer* errors, while paraphrase errors were the least prevalent. Among *intralingual errors*, the most prevalent was *ignorance of rule restrictions*, commonly noted in G8 (D/d), G4 (*Preposition + Noun*), G1, and G5 (*Adjective + Preposition*). *Interlingual errors*, specifically *L1 transfer*, were the primary cause of errors, especially in structures such as G8 (D/d) (*Verb + (Object) + Preposition + Object*) and G1 (*Noun + Preposition*).

Overall, the findings highlighted the substantial impact of *L1 transfer* and *intralingual* factors on learners' collocational difficulties regarding prevalent patterns. The total dominance of *intralingual* errors implies that learners are more susceptible to errors due to an insufficient or poor understanding of English grammar than to direct interference from their first language. However, learners' ability to accurately employ English collocations, including those frequently encountered in the language, was also influenced by negative transfer from Turkish. This emphasises the importance of instructional strategies tailored to enhance collocational competence and mitigate L1 interference in Turkish EFL learners. Additionally, the scarcity of *approximation* errors can be explained by the fact that writing allows students to plan, compose, and revise their work, which enables them to consult resources and deliberate on word choices, thereby reducing the likelihood of approximation errors. In addition, the written modality frequently requires a greater emphasis on language accuracy and formality, motivating learners to be more meticulous in their collocation selections (Harta et al., 2021).

All in all, the current study detected some problematic areas in the learners' use of grammatical collocations. It was evident that certain patterns were underused or overused by learners. Although the total frequency of deviant uses was not high, it was observed that learners had difficulties in using certain patterns which were considered to be

complex. This challenge could potentially limit the variety in their vocabulary and decrease the overall complexity of their communication by limiting their ability to employ a greater variety of collocations. Considering the study findings, it can be concluded that, even for proficient students, grammatical collocations stand as a problem in language learning. Therefore, special care should be given to the use of grammatical collocations rather than underestimating them.

5.2 Discussion

The current study is intended to provide an overall picture of the use of grammatical collocations in argumentative essays by Turkish ELT students. The findings are discussed in the sections below regarding the research questions. In section 5.2.1, the types and frequencies of grammatical collocations produced by Turkish ELT students and native speakers were discussed. Section 5.2.2 focuses on the frequency of grammatical collocational errors within the learner corpus. Finally, in section 5.2.3, sources of grammatical collocational errors within the learner corpus are examined.

5.2.1 The types and frequencies of grammatical collocations appropriately produced by Turkish ELT students and native speakers

The first research question aimed to find out the types and frequencies of grammatical collocations produced by Turkish ELT students and native speakers. The results indicated that grammatical collocations were used more frequently by native speakers than they were used by Turkish students. Regarding grammatical collocations in particular, the findings of this current study are in line with previous studies (e.g. Bahns and Eldaw, 1993; Biber et al., 1999; Wray and Perkins, 2000; Rankin and Schiftner, 2011; Özata; 2020, Sládková, 2022; Wang and Liu, 2022; Kim, et al., 2024)

This study's findings highlight the differences between learners' language systems and native speaker usage, especially in the context of grammatical collocations. In accordance with Wray and Perkins (2000), this study emphasises the difficulties encountered by L2 learners in utilising grammatical collocations relative to native speakers. Wray and Perkins (2000) highlighted that L2 learners frequently underutilised grammatical collocations, typically replacing them with simpler constructions due to inadequate collocational knowledge. This underutilisation indicates that learners struggle to internalise the formulaic characteristics of collocations, which represent a more

instinctive element of native language usage. Likewise, the findings of this study correspond with those of Kim et al. (2024), who examined the use of lexical and grammatical collocations among Korean EFL learners across different skill levels in comparison to native speakers. Utilising corpora from native speakers and Korean EFL learners, the study revealed that, irrespective of skill level, EFL learners employed significantly fewer collocations than native speakers. This finding highlighted the ongoing difficulties non-native speakers encounter in properly acquiring and utilising collocations, notwithstanding advancements in linguistic ability. According to Hunston (2003), native speakers find it easier to acquire formulaic language parts, such as collocations, due to their reliance on linguistic patterns, which reduces the amount of cognitive work involved in generating new sentences. In contrast, Hunston (2003) indicates that non-native speakers face increased challenges with these structures due to their often unpredictable usage, necessitating exposure and practice for mastery. This unpredictability may elucidate why non-native learners underutilised collocations such as G8(K) (*Verb + a possessive +V-ing*) and G8(M) (*Verb + Object + to be + Complement*) and chose more straightforward options (e.g. G8(F) (*Verb + bare infinitive*)) in their language use in this study.

The results of the current study indicated that specific structures were used more frequently in the native corpus than they were in the learner corpus (see Table 4.1, p. 61). To illustrate, the G8 (R) pattern (*It + Verb + Object + to Infinitive/that-clause*), an impersonal *'it'* construction, was significantly more frequent in the native corpus than in the learner corpus, reflecting a fundamental distinction between native and non-native speakers in sentence construction. Granger's (1998) research, which emphasised that learners frequently favour more uncomplicated, more straightforward sentence structures, is consistent with the findings of the current study. The researcher further asserted that learners tend to avoid impersonal constructions that are less intuitive and more idiomatic. Therefore, learners' underutilisation of these patterns in this study is likely due to their complexity, which necessitates advanced syntactic and pragmatic knowledge. This conclusion was further substantiated by Römer (2009), who observed that impersonal '*it*' constructions were a distinguishing feature of native academic writing, but they were much less common in learner output. It was further reported that these constructions frequently fulfil critical discourse functions, such as emphasis and hedging, which

learners may find challenging to replicate due to their lack of familiarity with these nuanced applications. In a similar vein, Huh and Hwang (2011) discovered substantial discrepancies in the selection of verbs and the omission of clauses between native and learner corpora. Despite the fact that native speakers employed a wide variety of verbs and contextualised these constructions, learners frequently resorted to simpler or less sophisticated forms, which highlighted the discrepancy in syntactic and stylistic proficiency. Larsson (2012) also emphasised the significance of anticipatory '*it*' constructions, observing their widespread use in native writing to convey subtlety, manage reader expectations, and facilitate effective hedging. Nevertheless, the pragmatic and stylistic functions of these patterns were frequently unfamiliar to L2 learners, which led to their restricted application. Consequently, as asserted by the relevant research, the results of the current study indicated that this underutilisation not only reflects a deficiency in syntactic knowledge but also underscores a more general obstacle to attaining academic sophistication and native-like fluency.

The pattern G8 (Q) (*Verb + (Object) + WH-clause/WH-phrase*) was also significantly higher in the native corpus due to its functional and syntactic versatility. This finding correlates to Boland et al.'s study (1995), which investigated *verb* argument structures and parsing in *WH-questions* and highlighted how native speakers process these constructions with high efficiency due to their familiarity with verb-specific argument patterns. This finding also aligns with that of Hunston (2003), who investigated the use of *WH-clauses* in verb complementation and highlighted their frequency in native English for purposes such as specification, inquiry, or meaning attribution. The study highlighted these formations' syntactic intricacy and semantic depth, which are frequently underutilised or misapplied in learner corpora. Additionally, Yoon (2008) attributed learners' underutilisation of embedded adverbial phrases and *WH-clauses* to the complexity of these structures. As a consequence, these studies collectively demonstrated that the G8(Q) pattern is indicative of good English usage, and its higher frequency in native corpora indicates a requirement for syntactic clarity and semantic precision - two qualities that are critical for advanced language ability. The relatively low frequency of this pattern in the learner corpus may indicate a deficiency in learners' linguistic competence, as it is instrumental in the production of precise and well-structured discourse. This absence may be the result of a preference for simpler, more familiar

structures, limited exposure to authentic usage, or challenges in mastering complex syntactic constructions.

The results also showed that the G6 pattern (*Adjective + to + Infinitive*) was considerably more prevalent in the native corpus than in the learner corpus. This outcome is consistent with Marco's (1999) examination of the *be + adjective + to + infinitive* structure, which combines grammatical and lexical components to convey intricate meanings, particularly in evaluative and modal contexts. The research by Marco (1999) emphasised that this construction serves as an example of both formulaic and productive language usage, allowing speakers to communicate nuanced attitudes, judgements, and assessments. It can be asserted that the increased frequency of G6 patterns in native speaker data indicates that their extensive exposure to diverse linguistic contexts enables them to employ these structures with greater ease and fluency. Conversely, learners may encounter difficulties in assimilating this pattern due to its syntactic complexity and the necessity for sophisticated linguistic processing.

The results also revealed that certain prepositional structures (G1 (*noun + preposition*), G5 (*Adjective + Preposition*), and G8(D) (*Verb + (Object) + Preposition + Object*) were higher in number in the native corpus than they were in the learner corpus. Among these prepositional structures, G8(D) (*Verb + (Object) + Preposition + Object*) was used significantly more often than in the native corpus. This is consistent with research showing that complex prepositional structures, such as G8(D) (*verb + object + preposition + object*), are more common in native corpora because of their idiomatic and syntactically integrated usage. To illustrate, Rankin and Schiftner (2011) asserted that native speakers utilise prepositional collocations more specifically and differently than learners, who frequently use prepositions interchangeably, leading to frequent overgeneralisation and errors. Nevertheless, it should also be noted that even if the frequency of these patterns was significantly higher in the native corpus, their occurrence, specifically G8(D) (*verb + object + preposition + object*), in the learner corpus was still noticeably substantial. As discovered by Gilquin and Granger (2011) the high frequency of this pattern may be indicative of learners' attempts to emulate the complexity of natives. Their research indicates that learners frequently utilise specific structures in an effort to demonstrate linguistic proficiency, particularly in academic or formal settings.

The analysis demonstrated that the G5 pattern (*Adjective + Preposition*), another prepositional structure, was more frequent in the native corpus than in the learner corpus, and is consistent with previous research (Bahns and Eldaw, 1993; Biber et al., 1999; Özata, 2020; Sládková, 2022; Wang and Liu, 2022). This finding highlights the challenges that learners face when trying to master *adjective + preposition* collocations, which are both idiomatic and prevalent in native English. Sládková (2022) provided additional evidence by conducting an analysis of English essays from Czech learners, which revealed frequent errors, such as the omission or misuse of dependent prepositions. According to the researcher, these errors were indicative of a developmental gap, as learners frequently encounter difficulty in internalising the fixed collocations that native speakers naturally employ. In the Turkish context, the findings of Özata's (2020) investigation into the use of collocations in student writings, specifically in the context of *adjective + preposition* patterns, revealed that native speakers employed substantially more collocations than Turkish L2 learners. It is indicated that this discrepancy emphasises learners' limited exposure and practice with such structures. Similarly, Wang and Liu (2022) conducted an analysis of argumentative essays by Chinese EFL learners and determined that the learners' inclination towards general adjectives and their avoidance of precise *adjective + preposition* constructions are a result of a limited lexical and syntactic range. The findings point to the obstacles that learners encounter in the process of mastering these collocations, as they reveal a substantial discrepancy between the utilisation of G5 (*Adjective + Preposition*) patterns by native speakers and learners. Bahns and Eldaw (1993) attributed this to the arbitrary nature of English prepositional selection as a key difficulty for learners, echoed by Biber et al. (1999), who noted that *adjective + preposition* combinations are common yet challenging for non-native speakers. It can be concluded that the unpredictability of English prepositional selection is a major factor in learners' difficulties acquiring G5 (*Adjective + Preposition*) patterns. This unpredictability, as noted by Bahns and Eldaw (1993) and Biber et al. (1999), makes such collocations particularly challenging to acquire. In contrast to rigorous grammatical rules, *adjective-preposition* collocations frequently employ idiomatic prepositions, which do not always adhere to logical patterns that learners can readily discern.

The research indicated that in addition to prepositional patterns, the G2 pattern (*Noun + to + Infinitive*) occurred significantly more often in the native corpus compared to the

learner corpus, highlighting a marked difference in usage between native and non-native speakers. This conclusion aligns with Bardovi-Harlig's (1997) assertion that learners frequently reduce or avoid intricate syntactic constructions such as *noun + infinitive* phrases because of the sophisticated cognitive processing they necessitate. Furthermore, Hinkel (2002) emphasised that learners often avoid infinitive formulations due to their intrinsic unpredictability and unexpected usage conventions, and this avoidance behaviour indicated that learners may prefer simpler and more known forms that are easier to generate and less syntactically challenging. Further insights into the functional significance of the G2 pattern were provided by Laso and John (2013), who conducted corpus-based analyses of collocational patterns in medical and academic English. Their research revealed the significance of *noun + infinitive* constructions in establishing logical and semantic relationships, particularly in specialised discourse domains. They argued that these structures are essential to producing idiomatic and fluent language, supported by the increased prevalence of such patterns in native data. As a consequence, the current study's results demonstrate the semantic specificity and discourse functions of G2 patterns in native corpora, which learners may underutilise as a result of their complexity and the cognitive demands associated with their use. This underutilisation may lead to a decrease in the production of cohesive and idiomatic language by non-native speakers.

It is also notable that both learners and native speakers tended to frequently use certain patterns: G8(E) (*verb + to + infinitive*) and G8(G) (*verb + V-ing*). This finding suggests that both learners and native speakers exhibit a strong tendency to use infinitive and gerund constructions frequently, emphasising their fundamental role in English grammatical structure. As Larsen-Freeman (2014) noted, the frequent use of infinitives and gerunds is indicative of their importance in English syntax and semantics. However, despite their prevalence in both corpora, these patterns were significantly higher in the native corpora, highlighting key differences in how native speakers and learners use these structures. This finding aligns with that of Deshors (2015), who examined constructions with verb complementation and emphasised that native speakers show a differentiated understanding of semantic and syntactic constraints when choosing between gerunds and infinitives. Learners, on the other hand, tend to be more rigid and less context-sensitive. Similarly, Taouis and Pérez (2019) analysed EFL learners' gerund and to-infinitive usage,

identifying that native speakers' intuitive grasp of these structures contributes to their significantly higher frequency and correctness in native corpora. Additionally, supporting this, Wallwork (2023) found that a wider frequency and range of verb-argument constructions containing these patterns can be found in native corpora, reflecting their integral role in fluent expression. To conclude, as previous research suggested, the disparity in the frequency and accuracy of gerund and infinitive constructions between native and learner corpora in this study can be attributed to a variety of key factors. Native speakers have a more sophisticated grasp of the syntactic and semantic rules regulating these structures, allowing for greater precision and flexibility. Conversely, learners frequently depend on strict, rule-based methods, which might result in errors or avoidance.

Besides frequently used patterns, the results also showed that learners overused certain patterns. As suggested by Altenberg and Granger (2001), learners tend to overuse specific clause structures to add detail, even if they do not use them precisely. This is consistent with the slight overuse of G8(F) (*verb + bare infinitive*) in the current study, which was significantly higher in the learner corpus. This finding also aligns with the findings of Tono (2004), Alqarni (2019), Taouis and Pérez (2019) and Sung (2022), who discovered that learners frequently employ this structure, particularly with modal verbs. This tendency was attributed to the high frequency of exposure to such constructions in instructional materials, textbooks, and everyday language input, making them easier to acquire and use. Additionally, it was reported that their structural simplicity reduces cognitive load, allowing learners to use these patterns more confidently. This was also endorsed by DeKeyser (2005), who emphasised that learners become more automatised by the recurrent use of frequent structures such as *verb + bare infinitive* in language input. The preference for patterns such as G8 (F) could be further reinforced by the fact that learners tend to employ simpler, high-frequency forms when constructing sentences, as indicated by studies on vocabulary frequency in learner corpora conducted by Cobb (2003) and Laufer (2003). As suggested by these studies, the over-utilisation of patterns such as G8(F) illustrates the impact of input frequency and simplicity on learners' linguistic output. Although dependence on high-frequency structures might enhance fluency and facilitate communication, it may also impede learners' capacity to vary their language use and correctly utilise more intricate formulations.

Another overused pattern in the learner corpus was G8 (L) (*Verb + that-clause*), which was significantly higher in that corpus. This is consistent with the assertion made by DeKeyser (2005), who emphasised the fact that learners at intermediate levels frequently utilise *that-clauses* to communicate nuanced concepts, thereby mitigating limitations in their vocabulary or syntactic variety. The findings of Gilquin et al. (2007), which observed that learners rely on *that-clauses* as a straightforward method to express complex ideas, can be supported by the common use of *verb + that-clause* among learners. The structure's simplicity compared to more complex sentence connectors may be the reason for the excessive reliance on *that-clauses,* as research highlights that this overuse stems from non-native speakers' reliance on simpler syntactic structures due to limited exposure to varied sentence constructions. Calude (2017) provided further evidence for the overproduction of clause complements by non-native speakers. As a result, learners' frequent usage of that-clauses in this study highlights their reliance on more recognisable and structurally simpler patterns to convey complicated ideas. This tendency implies that a lack of exposure to a variety of syntactic structures and an excessive dependence on formulaic formulations may hamper learners' capacity to generate more complex and varied sentence forms.

According to Selinker's (1992) concept of 'interlanguage', learners tend to use familiar structures, even in an inaccurate manner, as a consequence of the limited options available within their developing language system. Thus, the slight overuse of *verb + adverbial* structures, specifically G8(P) (*Verb + (O) + Adverbial*) by learners, might be attributed to the scarcity of alternative expressions in the interlanguage. This finding aligns with Kim et al. (2024) who established that when particular patterns are concerned, namely, G8(P), *verb + adverbial*, non-natives use this pattern more frequently than native speakers. This finding may be attributed to the fact that L2 learners have a tendency to prefer particular combinations of high-frequency collocations and overuse them (Granger, 1998; Durrant and Schmitt, 2009; Salido et al., 2018). As a consequence, as indicated by research, the widespread usage of *verb + adverbial* structures by students raises the possibility that interlanguage gaps may be related to their over-reliance on these patterns. Although these structures facilitate communication, they also indicate a deficiency in the diversity of learners' grammatical repertoire. This overreliance signifies

that learners favour familiar and frequent patterns instead of less common but perhaps more contextually suitable options.

In addition to high-frequency patterns, low-frequency patterns were also observed in both corpora. The low frequency of G8 (K) (*Verb + a possessive + V-ing*) and G8 (M) (*Verb + Object + to be + Complement*) aligns with previous studies that emphasise the complexity and specialised nature of these patterns, which makes them less common in general language use (Bardovi-Harlig, 2002; Collins, 2004; Hassan, 2014; Deshors and Gries, 2016). Bardovi-Harlig (2002) underscored that the inherent complexity of complex syntactic patterns, such as *possessive + gerund* forms, necessitates their acquisition at advanced stages of language proficiency. Learners frequently lack a nuanced comprehension of grammatical relationships, particularly at lower proficiency levels, which is necessary to comprehend these patterns. Deshors and Gries (2016) further underlined that learners find it particularly challenging to form *verb + possessive + V-ing* and other structures that are constrained by certain lexical and grammatical dependencies. In brief, the low frequency of the patterns G8(K) and G8(M) in both native and learner corpora may be attributed to these pattern's specialised syntactic and semantic roles. These findings underscore the contextual and lexical constraints of G8(K) and G8(M), which learners might find challenging due to their infrequency and complexity in natural language use.

In summary, as Cobb (2003) observed, even advanced English learners often struggle to fully grasp the breadth and nuance of English phraseology. Non-native writers tend to rely heavily on high-frequency collocations, often neglecting less common but strongly associated collocational patterns—linguistic elements that are more salient and readily accessible to native speakers (Durrant and Schmitt, 2009). This tendency reflects a reliance on predictable and formulaic language, which may contribute to the lack of idiomatic richness typically found in native-speaker writing. Such findings align with usage-based models of language acquisition, which posit that linguistic competence is shaped by the frequency and variability of input. These models help explain the observable gap in learners' ability to produce diverse and idiomatic expressions, underscoring the challenges non-native speakers face in achieving native-like fluency and naturalness in productive language use (Ellis, 2002).

5.2.2 The frequency of grammatical collocational errors within the learner corpus

The second research question was to ascertain the frequency of grammatical collocational errors in the learner corpus. In the current study, the results indicated that the overall deviation rate across all patterns was 3.91% (see Table 4.4, p. 66). The relatively low frequency of errors within the learner corpus can be attributed to the fact that the participants in the current study were intermediate-level ELT students since there exists ample evidence of the relationship between learners' overall English proficiency and their collocational performance, with numerous studies emphasising the importance of collocational competence for language proficiency. This relationship implies that learners' competence in using collocations, which are crucial for acquiring fluency and authenticity in a second language, increases in parallel with their general language proficiency (e.g., Bahns, 1993; Keshavarz and Salimi, 2007; Chen, 2008; Du et al., 2022; Masrai, 2023).

Nevertheless, despite their moderate level of proficiency in English, the findings revealed that the participants encountered difficulties with some complex structures. To illustrate, the deviation rate of 39,29% of the use of pattern G8(R) (*It + Verb + Object + to Infinitive/that-clause*) underscores the challenges learners encounter when employing impersonal terms, such as "*It annoys me that she never….*". This aligns with the findings of Chen (2024), who conducted an analysis of the English writing of high school students and identified a significant number of errors in verb usage and collocations, particularly in syntactic constructions such as G8(R). This may be ascribed to the fact that idiomatic and formulaic expressions, such as impersonal constructions, are particularly difficult for learners to acquire because they must be picked up as fixed units (Alhaysony, 2017). In addition to this, in academic and professional writing, impersonal patterns are prevalent, as noted by Biber et al. (1999) and Conrad and Biber (2001), in contrast to their rarity in conversational English. Consequently, this discrepancy generates additional challenges for students who are more familiar with informal registers, a difficulty further compounded by the inadequate emphasis on these patterns in traditional ESL curricula (Richards, 2002). Collectively, these results demonstrate the diverse obstacles that impersonal constructions present to language learners due to their complex structure.

However, the patterns with the highest frequency of errors in this study are inconsistent with the findings of Aye (2021), who investigated the grammatical collocation errors in narrative essays of a total of 45 Myanmarese postgraduate university students specialising in English. Contrary to the findings of the current study, the highest frequency of errors was reported in pattern G8 (D/d) (*verb + preposition + object (or) verb + object + preposition + object*). This disparity may be attributed to the differences in the scope, focus and the first language backgrounds of participants in the two studies. The results of the current research may have been influenced by the fact that Aye's (2021) study analysed a smaller dataset (45 essays) than the current study (206 essays). Furthermore, the essay genres were distinct: Aye's (2010) investigation concentrated on narrative essays, whereas the present investigation investigated argumentative essays. Additionally, participants' L1 were different; in Aye's study, participants' L1 was Burmese, while in this study, participants' L1 was Turkish. Additionally, the participants in Aye's study were postgraduate students with a higher level of proficiency than those in the present study. Hence, the contrasting patterns of grammatical collocation errors observed in the two studies may be attributed to the variations in sample size, genre, participants' first language backgrounds and their proficiency level.

The structure G8(K) (*Verb + possessive + V-ing)* also demonstrated notable deviation rates, with 25% of instances in the learner corpus. Supporting this, Can (2017) discovered that *verb + V-ing* constructions were among the least error-prone patterns in an analysis of verb errors among Turkish EFL learners. Inaccuracies were frequently observed in more intricate contexts that necessitated the complex use of gerunds or participles. This can be attributed to the fact that possessive gerund formulations are generally mastered solely at advanced competency levels due to the requirement for a nuanced comprehension of both possessive forms and gerunds, as stated by Bardovi-Harlig (2002). Larsen-Freeman (2014) further noted that learners frequently encounter difficulties with this pattern because of its infrequent occurrence in natural language exposure and formal education. Numerous English textbooks and curricula stress prevalent verb-gerund pairings, therefore, neglecting possessive gerund constructions. Consequently, as suggested by previous research, the elevated deviation rates of this structure in this study could be attributed to the fact that learners struggle with recognising the appropriate verb-complement interactions, leading to frequent deviations from native-like usage. A key

factor contributing to these errors is the complexity of verb-object-complement constructions, which necessitate an advanced comprehension of verb function in various syntactic contexts. G8(M) structures demand nuanced judgment, which learners may find difficult, especially in productive language tasks like speaking and writing, in contrast to simpler verb patterns that adhere to more predictable principles. Furthermore, because these constructions can have multiple meanings depending on the context, students frequently find it difficult to apply them effectively.

G8(M) (*Verb + Object + to be + complement*) pattern also demonstrated notable deviation rates, with 25% of instances in the learner corpus. This finding is consistent with that of Bourke (2007), who highlighted that errors in verbal complementation are frequent among ESL learners. The researcher emphasised the challenges learners face due to the subtle distinctions in verb patterns and the limited effectiveness of formal rules in teaching these structures. This can also be ascribed to the observation by Larsen-Freeman (2014), who highlighted that verb-object-complement structures are challenging to master because of their flexibility and dependence on context. Students frequently simplify or overgeneralise these structures, resulting in inaccuracies. Therefore, mastering verb-object-complement structures may require a deep understanding of syntax, verb behaviour, and contextual appropriateness.

The use of the pattern G8 (J) (*Verb + Object + V-ing*) exhibited a 23.08% deviation rate, signifying considerable learner challenges with gerund usage after verbs with objects. This finding aligns with Granger and Thewissen's (2005) examination of corpora containing error markers, highlighting that constructions with gerundive complements are frequently misused by non-native speakers. Their work suggests that such errors are influenced by learners' tendency to simplify complex structures, which reduces the frequency of correct use of the G8(J) pattern. In the Turkish context, Can (2017) analysed EFL learners' verb errors, noting that wrong verb forms, including errors in gerund constructions, are among the most frequent. The findings highlighted that learners often fail to identify contexts requiring *verb + object + v-ing*, leading to deviations from native norms. The elevated frequency is attributable to Larsen-Freeman's (2014) finding that learners frequently find it challenging to distinguish between gerund and infinitive forms because of the absence of explicit, definitive rules regulating their application.

G8 (I) (*Verb* + *Object* + *Infinitive*) exhibited a 14.40% deviation rate, indicating possible challenges for learners, presumably attributable to the intricate rules regulating infinitive usage. As noted by Ellis (2006), such errors may stem from inadequate exposure to the target structure, leading learners to overgeneralise from simpler sentences devoid of objects. Taouis and Pérez (2019) also discovered the difficulties learners face with infinitive complements, particularly in constructions requiring *verb + object + infinitive* and emphasised the complexity of this structure for L2 learners. Their examination of Spanish learners' writing demonstrated recurrent errors in these structures, highlighting the complex nature of this pattern for L2 learners. Consequently, the research by Taouis and Pérez (2019) concluded that the intricacy of these structures frequently surpasses learners' grammatical proficiency, resulting in ongoing difficulties in their appropriate usage.

The pattern G7 (*adjective + that-clause*) also exhibited a relatively high rate of 11.24%. This finding aligns with that of Altenberg and Granger (2001), who observed that learners often excessively employ that-clauses after adjectives, typically resulting from a restricted understanding of sentence structure variation. Huh and Hwang (2011) further investigated similar errors in the academic writing of Korean learners, discovering patterns of misuse, including redundant or omitted that-clauses. These errors were ascribed to learners' dependence on informal or simplified grammatical structures, particularly in academic contexts. Kochmar and Briscoe (2015) provided additional insights highlighting that learners frequently choose unsuitable adjectives or neglect crucial *that-clauses* owing to inadequate lexical and syntactic expertise. Interestingly, the findings of the current study contrast with those of Alsulayyi (2015), who examined grammatical collocations, specifically the patterns G1-G7, among 10 Saudi English learners. Alsulayyi documented reduced error rates for the *adjective + that-clause* construction. The discrepancy may be ascribed to the differing contexts of the studies and the fact that the data in Alsulayyi's study was drawn from a comparatively smaller number of articles than in the current study. This highlights the importance of considering learner background, corpus composition, and task type when analysing grammatical collocation patterns. In conclusion, the findings of this study reinforce the notion that learners tended to overuse *adjective + that-clauses* due to a restricted understanding of syntactic variation, as highlighted by Altenberg and Granger (2001). The patterns of

erroneous use identified by Huh and Hwang (2011) and Kochmar and Briscoe (2015) further suggest that this issue may not merely be a result of frequency but also stems from learners' limited lexical and syntactic awareness, which leads to redundant, omitted, or inappropriate *that-clauses* in their writing.

With a moderate (10.95%) deviation rate, learners also seem to encounter difficulties with G8 (H) (*Verb + Object + to Infinitive*). This finding aligns with that of Kim and Yoo (2015), who analysed the use of to-infinitives by Korean EFL learners and found frequent errors in verb complements. The study attributed these errors to the fact that the learners were not exposed to varied input. As previous research further suggested, errors may arise as learners struggle to recall which verbs necessitate an infinitive following the object. Larsen-Freeman (2014) asserted that understanding infinitive complements poses challenges due to the diversity in verb requirements. The learner's task is further complicated by the fact that the proper use of these complements frequently depends on the meaning and intention of the verb. For these rules to be used effectively, learners must be able to memorise specific verb patterns and understand how they are used in different circumstances, both of which can be challenging.

The deviation rate of 10.26% for pattern G8 (C) (*Verb + Direct Object + for + Indirect Object*) indicates that learners had difficulties using this pattern accurately. Studies related to this structure showed that learners struggle to accurately place the direct or indirect object. For example, Millaku (2015) examined the use of direct and indirect objects in Albanian and English, highlighting that learners find this pattern challenging and misplace or omit indirect objects in English, especially in constructions such as *Verb + Direct Object + for + Indirect Object*. Similarly, Taouis and Pérez (2019) examined noun and direct object collocations in the written English of Spanish learners, uncovering errors in the application of verb patterns with indirect objects. Students often encountered difficulties assigning suitable responsibilities to objects, resulting in deviations from natural standards. Consequently, the high frequency of deviation in this pattern can be ascribed to the fact that learners find this structure challenging due to difficulties in accurately assigning indirect objects in phrases maybe owing to insufficient exposure to the subtleties of these prepositions in formal education.

The pattern G8 (N) (*Verb + Object + Complement*) also had a moderate deviation rate (8,15%). This aligns with research by Taouis and Pérez (2019), who examined Spanish learners' collocational usage, particularly in constructions involving direct objects and complements. Chen (2024) presented additional evidence of this issue through a corpus-based investigation of English writing errors. The research highlighted that verb-complement patterns typically lead to learner errors. This can be attributed to the fact that L2 learners find recognising and utilising suitable complements challenging (Swan, 2005). As Taouis and Pérez (2019) further suggested, the frequent misuse of verb complements can also be ascribed to insufficient emphasis on collocational teaching. The research posited that conventional grammar-focused instruction fails to sufficiently tackle collocational patterns, resulting in learners being uninformed of the appropriate verb complement patterns.

Regarding grammatical collocations involving prepositions, there was a relatively moderate frequency of errors. This can be considered consistent with the studies showing that errors related to the use of prepositions are common among grammatical errors (e.g. Chen, 2008; Demirel, 2017; Kırmızı and Karcı, 2017; Taşçı and Ataç, 2018; Alsulayyi, 2015; Khorshid and Alishah, 2021). To illustrate, in this study, a deviation rate of 5.77% in G1 (*noun + preposition*) showed that learners had moderate difficulties with prepositional structures. This discovery is consistent with prior research (e.g., Chen, 2008; Alsulayyi, 2015; Khorshid and Alishah, 2021), which also identified *noun + preposition* as a difficult pattern. It is notable that these studies did not analyse all the patterns proposed by Benson et al. (2010), contrary to the current study. However, it was consistently reported that *noun + preposition* collocations had the highest level of deviance. In the same vein, the relatively moderate deviation rate observed for this pattern in the current study corroborates and reinforces their findings. Nevertheless, this finding does not align with that of the aforementioned study by Aye (2021). The researcher found that *noun + preposition* combinations had one of the lowest deviation rates. This discrepancy may also be attributed to the variations in sample size and genre, as mentioned above. This implies that learners' challenges with grammatical collocations could vary depending on the context and be impacted by the types of material they are exposed to as well as the teaching strategies they encounter. However, given its moderate deviation rate in this study and its recognition as a troublesome pattern in previous

research, it can be concluded that learners frequently encounter difficulties with the *noun + preposition* structure. The consistent results from several research indicate that learners have difficulty choosing and placing prepositions in *noun + preposition* collocations. This difficulty is probably caused by the idiomatic and non-transparent nature of these constructions in English.

Among prepositional patterns, the pattern G8 (D/d) (*verb + preposition + object (or) verb + object + preposition + object*) showed a 5,18% deviation rate in total, which indicates moderate difficulties with prepositional collocations. This is in line with previous research by Chen (2018) and Özata (2020). Chen's (2018) study identified *verb + preposition* errors as one of the most prevalent collocational issues. In the same vein, Özata (2020) discovered that even advanced-level learners exhibited a high frequency of errors with this pattern, underscoring the persistent challenges these structures pose. According to Demirezen (2021), the difficulty that Turkish learners of English frequently encounter with these combinations primarily stems from the absence of analogous structures in Turkish morphology and syntax. Prepositional verbs are non-existent in Turkish, which complicates the comprehension of their function in English for learners. Furthermore, *verb + preposition* structures are frequently confused with *verb + particles*, which serve a distinct grammatical function. Particles, as stated by DeCapua (2008), are prepositions and adverbs that are combined with verbs to create new verbs, frequently losing their original prepositional or adverbial meaning. Thus, the learners' comprehension is further complicated by the complexity of distinguishing between prepositional verbs and phrasal verbs, as certain verbs can serve as either depending on the context. This dual functionality leads to misunderstanding because phrasal verbs frequently have idiomatic meanings and may or may not require an object, whereas prepositional verbs require an object to fulfil their meaning.

Additionally, the pattern G5 (*Adjective + Preposition*) showed a 4,15% deviation rate in total, among prepositional patterns, indicating moderate difficulties with prepositional collocations. This is consistent with several studies which found that *adjective + preposition* errors were among the common ones (e.g., Saud, 2018; Sari and Gulö, 2019; Paris et al., 2023). The research underscores the pervasive difficulties encountered by EFL learners regarding *adjective + preposition* collocations. Saud (2018) identified these as the predominant grammatical collocation errors among learners, while Sari and Gulö

(2019) noted a significant prevalence of such errors in essays by first-year Indonesian English Literature students, corroborating challenges in selecting suitable prepositions. Likewise, Özata (2020) discovered that both pre-intermediate and advanced learners struggled with these structures, signifying enduring issues across proficiency levels. Paris et al. (2023) additionally reported frequent errors in *adjective + preposition* combinations among Asian EFL learners, highlighting the global scope of these challenges. Conversely, Khorshid and Alishah (2021) discovered that Iranian EFL learners experienced little challenges with this pattern. This variance may indicate variations in learners' linguistic backgrounds, pedagogical methods, or exposure to English.

All in all, the errors in prepositional structures might be attributed to the fact that English prepositions are frequently a challenge for even the most advanced learners, as the choices made are largely arbitrary and necessitate extensive practice to master (Liao and Fukuya, 2004). Furthermore, as stated by Šebestová (2021), the function of prepositions in text organisation and their analytical nature in English frequently present difficulties for learners because of limited exposure and disparate linguistic frameworks.

The results also revealed that certain patterns showed lower deviation rates. To illustrate, pattern G6 (*Adjective + to + Infinitive*) also had a low deviance rate (2,02%), which indicated that learners demonstrated proficiency in this structure, frequently employing it to convey ability or emotion. Nation (2001) ascribed the simplicity of this structure to its prevalent use in conversational English and its uncomplicated formulation, enabling learners to attain familiarity rapidly. This finding, however, contradicts that of Huriyah (2014), who examined the use of gerunds and to-infinitives among second-grade students in Jakarta. The researcher indicated a recurrent error between gerunds and to-infinitives, adversely affecting learners' precision in adjective complementation. This discrepancy may be explained by variations in the proficiency levels of participants in the two investigations. The present study concentrated on learners with a higher degree of English competence, which may explain their increased acquaintance with the *adjective + to-infinitive* structure. Additionally, the varying circumstances and educational environments of the two investigations may have impacted the outcomes.

The lowest deviation rate (0.63%) was observed in pattern G8 (F) (*Verb + bare infinitive*). This was corroborated by Alqarni (2019), who employed corpus pattern analysis to

investigate verb usage and observed that learners are proficient at generating bare infinitives. As noted by DeKeyser (2005), learners are more likely to acquire structures with clear, consistent norms, particularly when they are reinforced through recurrent exposure. Bare infinitives are frequently encountered in conjunction with modal verbs and causal forms, which renders them highly accessible and simpler for learners to internalise. Consequently, as supported by the previous research, learners' robust comprehension of bare infinitives is illustrated by the low deviation rate for G8(F) and can be ascribed to its uncomplicated structure.

In conclusion, the research indicates that learners possess a comprehensive comprehension of more basic structures, such as *verb + bare infinitive* and *verb + v-ing*, which are more straightforward and frequently employed with modal verbs and causal forms. Nevertheless, they encounter difficulties when confronted with more intricate structures such as G8(K) (*Verb + possessive + V-ing),* G8(M) (*Verb + Object + to be + complement*) and G8(R) (*It + Verb + Object + to Infinitive/that-clause*) because of their structural flexibility and contextual variability, which can result in overgeneralisation or simplification errors. Learners also encounter substantial obstacles when dealing with prepositions, which are frequently arbitrary and necessitate extensive practice. The analytical nature of English prepositions and their role in text organisation also present additional challenges, particularly for learners with limited exposure to authentic language use or whose native languages differ considerably in prepositional systems. The results highlight the necessity of personalised instruction and practice to assist learners in mastering these more intricate components of English grammar.

5.2.3 The sources of the grammatical collocational errors within the learner corpus

The third research question aimed to identify the sources of the grammatical collocational errors within the learner corpus. The overall findings indicated that there was a prevalence of *intralingual* errors over *interlingual* errors (See Table 4.5, p. 68), occurring as a result of "faulty or partial" learning of the English language (Richards et. al, 2002). This finding is consistent with studies which show that intralingual errors occur more frequently than errors due to interference from the first language (L1) when learners use grammatical collocations (e.g. Özışık, 2014; Paris et al., 2023; Park, 2024). This suggests that learners are more susceptible to errors as a consequence of an inadequate or deficient

comprehension of English grammar than to direct interference from their first language (L1). As a result, the writings of EFL learners frequently contain *intralingual errors*, which are routinely analysed to determine the strategies that EFL students employ in their writing (Jassim, 2020). These errors are likely the result of overgeneralisation, the incorrect application of rules, and false assumptions about grammatical structures, highlighting the necessity of targeted instruction that addresses learners' misconceptions and encourages the precise use of grammatical collocations.

Within *intralingual errors*, with 143 occurrences, *ignorance of rule restriction* was the predominant error category. A common category of intralingual errors in second language acquisition was *ignorance of rule restriction*, which represents learners' incorrect application of grammatical rules due to their insufficient comprehension and where learners often misinterpret the criteria governing rule applicability (Römer, 2009). This frequently results from a lack of exposure to various linguistic situations, which causes incorrect constructions. For example, students may say, "He made me to go," rather than the proper phrase, "He made me go" (Amara, 2015). The next frequent source of error was *incomplete application of rules*, indicating a challenge in the comprehensive application of rules across diverse situations. This suggests that learners frequently comprehend a rule yet fail to apply it comprehensively, consistent with Brown's (2000) observations regarding incomplete procedural knowledge. This finding also suggests that learners often understand basic rules but struggle with full application due to a lack of varied context exposure (Ellis, 2006). A high occurrence of *false concept hypothesised* errors was observed in 113 instances, indicating a widespread misunderstanding of specific grammatical rules among learners (Alan, 2015). These errors typically arise when learners form incorrect hypotheses about the target language, often due to incomplete or misleading input. Al-Tamimi (2006) characterises this type of error as stemming from insufficient gradation in instructional materials, whereby the progression of grammatical structures is not adequately sequenced or scaffolded. As a result, learners may infer inaccurate rules based on limited exposure. Such errors highlight the critical role of carefully structured input and instructional clarity in second language acquisition, particularly in ensuring that learners form accurate mental representations of grammatical systems.

The next error type within *intralingual errors* was *overgeneralisation,* with 62 instances occurring when a learner uses a different language structure instead of adhering to the proper usage of one (Heydari, 2012). This error type demonstrates learners' excessive application of known norms to unsuitable circumstances (Lightbown and Spada, 2013). According to Olasehinde (2020), the primary reasons for *overgeneralisation* and false concept hypothesised are neglecting exceptions, regularising language, and simplifying language, assuming that all structures adhere to the same patterns. As a result, these findings indicate that learners often misinterpret or overgeneralise grammatical rules, applying them to contexts where they are not appropriate as they endeavour to adapt the principles they have acquired in one context to another without a comprehensive understanding and fail to identify irregularities in grammatical rules. Learners frequently recognise recurring grammatical structures and presume that they adhere to universal rules. They may fail to recognise linguistic patterns that regulate irregular forms, as they may force rules onto exceptions due to their expectation of consistency. Although this process is a critical component of learning, it can result in systematic errors that persist if not explicitly rectified.

Interlingual Transfer errors associated with *L1 transfer* manifested prominently across patterns, with 234 instances. These errors indicate a dependence on native language syntax, resulting in inappropriate English sentences (Odlin, 1989; Selinker, 1992). Errors in *L1 transfer* indicate a tendency to impose native patterns onto English, aligning with Selinker's (1972, 1992) interlanguage theory. This may also indicate the instructional setting in which learners were instructed to minimise dependence on their native language. Numerous research papers offer evidence indicating that L1 interference significantly impacts L2 learners' collocation performance. This frequently leads to inaccurate assumptions, such as the assumption that the order of English words or the use of prepositions is identical to that of the native language (e.g. Bahns, 1993; Fan, 2009; Nesselhauf, 2003; Mahmoud, 2005; Taouis and Pérez, 2019; Yang et.al, 2020; Kayaer, 2023). Specifically, in the Turkish context, Kayaer (2023) focused on the errors of ELT preparatory students in using English grammatical collocations, particularly prepositions. The findings showed that participants make errors even when using grammatical collocations that are frequently used in the target language and that participants' performance is similarly impacted by negative transfer of L1 (Turkish). This implies that

being acquainted with frequent structures does not necessarily prevent errors. This suggests that learners may not have fully internalised these collocations as fixed units, resulting in the incorrect construction of phrases despite their exposure to them. These studies collectively emphasised that learners' capacity to accurately employ English collocations, including those that are frequently encountered in the language, is substantially impacted by negative transfer from Turkish. This underscores the necessity of instructional strategies that are specifically designed to address L1 interference and improve collocational competence among Turkish EFL learners.

Communicative Strategies were infrequently used by learners throughout patterns. Within the category of paraphrasing, there existed merely 15 occurrences of *approximation*, denoting the utilisation of incorrect vocabulary items or structures (Dravishi et al., 2011). This restricted usage indicates that learners can favour a stringent, accuracy-oriented method of language creation instead of utilising more adaptable, communicative techniques to address deficiencies in their knowledge. *Approximation* errors, which occur when learners substitute an imprecise term for a precise one, are relatively uncommon in written essays that address collocation use (Do et al., 2023). This rarity can be attributed to the fact that writing provides learners with ample time to plan, compose, and revise their work, allowing them to consult resources and contemplate word choices, thereby decreasing the probability of approximation errors. Additionally, the written modality frequently necessitates a larger degree of attention to formality and language accuracy, which in turn encourages learners to be more meticulous in their collocation selections (Harta et al., 2021). Written texts, particularly academic or professional writing, necessitate formality, which encourages learners to use precise expressions rather than approximations. Learners are motivated to employ well-established collocational patterns due to the necessity of coherent sentence structure and logical progression. These factors collectively contribute to the decreased frequency of approximation errors in written essays that are associated with the use of collocations.

In the following sections, the sources of errors, namely *L1 transfer, false concept hypothesised, overgeneralisation, incomplete application of rules, ignorance of rule restriction*, and *approximation,* in each grammatical pattern are examined in detail. By using sample sentences from the learner corpus, both the proportion of errors within each pattern and their total rate across all patterns are discussed in each section.

5.2.3.1 L1 transfer errors

In overall terms, the total number of *L1 transfer* errors indicated that this source of error was mostly observed in G8(D)/(d) (*Verb + (Object) + Preposition + Object*) with a percentage of 40,60% (95 cases out of 234). G1(*noun + preposition*) was the second highest (11,97%; 28 out 234) across all patterns. On the other hand, the lowest percentages were seen in patterns G8(H) (*Verb + Object + to Infinitive*) (3 out of 234; 1,28%) and G8(R) (*It+ Verb + Object + to Infinitive/that-clause*) (2 out of 234, 0,85%) (See Table 4.7, p. 70). This indicates that L1 interference was distributed across different patterns rather than affecting all structures equally, confirming that while L1 influence was an issue in some areas, it was not equally problematic in all grammatical patterns.

When *L1 transfer* errors were examined regarding the frequency of this error within each pattern, the results revealed that *L1 transfer* error was dominant in prepositional structures in the learner corpus, specifically G8(D)/(d) (*Verb + (Object) + Preposition + Object*), G4 (*preposition + noun*), G5 (*adjective + preposition)* and G1(*noun + preposition*) (See Table 4.10, p. 74). As noted by Swan and Smith (2001), prepositions are especially difficult for Turkish learners to understand since Turkish uses case suffixes rather than distinct prepositions, which frequently results in mistranslations. For learners, the most problematic prepositional pattern regarding *L1 transfer* was G8(D)/(d) (*Verb + (Object) + Preposition + Object*). *L1 transfer* was the most frequent type of error within this pattern, with a percentage of 51,63% (95 out of 184). This indicates that incorrect verb-preposition-object combinations in English seem to be the result of learners transferring prepositional structures from their native language (L1), in which prepositions may be unnecessary or utilised differently. For instance, in the learner's sentence, "*They like to listen music while studying to concentrate.",* '*listen*' was employed without the word '*to*' as a result of Turkish syntax, which frequently does not require prepositions such as '*to*' in equivalent expressions. The direct object of the verb in Turkish is followed by the verb without any additional indicators (e.g., '*müzik dinlemek'* translates to '*listen to music*'). As Demirezen (2021) emphasised Turkish learners of English frequently encounter difficulties with *verb + preposition* combinations, primarily as a result of the absence of analogous structures in Turkish morphology and syntax. Prepositional verbs are nonexistent in Turkish, which complicates the comprehension of their function in English for learners. This finding is also in line with previous research by Aye (2021), who

highlighted that in the use of this structure, the errors in the writing of the Myanmarese graduate students were the result of the intrusion of their native language. However, this study's results are in contrast to those of Özışık (2014), who discovered that Turkish EFL students made more intralingual than interlingual mistakes while using prepositions. 30 upper-intermediate students took a controlled 60-sentence gap-filling examination as part of Özışık's study, which examined their native language influence and prepositional competency. Methodological discrepancies could be the cause of the disparity: 179 free-writing essays were examined in this study, whereas Özışık employed a methodical gap-filling strategy. While gap-filling exercises centre emphasis on particular grammatical rules, potentially highlighting intralingual errors, free writing allows for the use of natural language, making learners more prone to errors influenced by their native language. The task's complexity could be another factor. While gap-filling activities lessen cognitive strain and increase intralingual errors, argumentative essays require greater cognitive effort and rely more on L1 structures (Wang et.al, 2024).

L1 transfer error was also prevalent in G1(*noun + preposition*), accounting for 28 instances out of 65 (43,08%), representing the predominant source of error in this structure and in overall terms. The rate was the second highest across all patterns (11,97%), indicating the substantial impact of first language (L1) interference on the production of *noun + preposition* structures. This implies that learners may have encountered major difficulties in accurately selecting the appropriate preposition when constructing noun phrases, which is likely attributable to the structural differences between their L1 and the target language. For instance, in the following sentence from the learner corpora, "*They will have difficulties <u>at</u> telephone services…*" learners may have habitually used '*at*' in their native language, whereas English necessitates '*with*'. The learner may have translated the prepositional usage directly from Turkish to English. In Turkish, context and case suffixes (e.g., -de, -den) establish links between nouns and their complements, resulting in problems when translating these structures into English prepositions. This finding is in line with that of Sravanan and Lakshmikanthan (2020), who investigated the challenges in the use of preposition combinations and analysed the sources of errors in the essays written by 60 Engineering students whose L1 was Hindu. The results showed that errors in the use of prepositions after nouns, verbs, and adjectives were due to L1 interference and ignorance of rules. This indicates that learners err with

noun-preposition combinations, driven by the disparities in prepositional usage between English and their native language.

The findings showed that for pattern G5 (*adjective + preposition), L1 transfer* errors occurred 16 times out of 30 cases (53,33%) within the specific pattern, while in overall terms, the rate of this error type was 6,84% across all patterns (out of 234). This shows that although it was responsible for a minor amount of total errors across other patterns, *L1 transfer* was the most common error type in this particular pattern. The dominance of *L1 transfer* indicates that learners seem to excessively depend on their L1 due to insufficient proficiency in the L2. In the given sentence, '*The additional information was irrelevant with solving the current issue…',* the word '*with*' was used incorrectly in place of '*to*' likely due to direct translation from Turkish. In Turkish, learners are influenced to use the Turkish preposition equivalent '*ile*' (with) rather than the proper English preposition '*to*' by phrases such as '*konuyla alakasız*' (*irrelevant with the issue*). The dominance of L1 transfer regarding this pattern is consistent with research by Taşçı and Ataç (2018), and Ataç and Önal (2022) showed how L1 interference significantly affects *adjective-prepositional* phrases. The crucial significance of the first language in learning these collocations was further underscored by Yuvayapan and Yükselir (2021), who noted that learners' collocation errors are frequently attributed to L1 interference.

In another prepositional pattern, G4 (*preposition + noun*), there was a significant prevalence of errors due to *L1 transfer* with 14 instances (31,11%), the second most dominant type of error, with an overall rate of 5,98%. Although *L1 transfer* seems to be a major problem in this particular pattern, the overall rate suggests that learners also have trouble using prepositions in other patterns. Learners are likely to find it challenging to use prepositions correctly before nouns, as evidenced by native language interference accounted for over one-third of the errors in this pattern. This implies that errors were caused by improper collocations, interference from direct translations, and prepositional choices based on L1. This is in line with the findings of Yamashita and Jiang (2010), Sanjaya and Bram (2020) and Neumanová (2023), which highlighted the fact that prepositions are particularly challenging for EFL students and many errors within this structure were attributed to L1 interference. Supporting this, in this study, it was observed that students frequently depended on the prepositional frameworks of their original language, leading to mistakes such as '*with this way*' instead of '*in this way*' in the

following sentence, "*With this way I think the problem can be solved.*" This error might arise since in Turkish, '*bu yol ile*' means '*with this way*', and students misused the Turkish preposition equivalent '*with*' in English, where '*in*' is necessary. As asserted by Kennedy (2003), since different prepositional combinations are used with the same lexical word, which has several meanings that are frequently unclear, learners find prepositions challenging to employ and make errors.

Apart from prepositional patterns, *L1 transfer* errors had the highest frequency within several patterns (e.g. G3, G8(A), G8(B), G8(N), G8(O), G8(Q), G8(S)). The results indicated that among them, in G8(A) (*Verb + DO + to + I O = V + IO + D O), L1 transfer* errors had the highest rate (82,35%), with 14 instances out of 17, underscoring the impact of learners' native language syntax on the pattern. In overall terms, the rate was 5,98%. This shows that while it did not represent a major portion of errors across all grammatical structures, *L1 transfer* was a prevalent issue within this specific pattern. This was noticeable in the placement of prepositions and the word order of direct and indirect objects. In the given example, "*Teachers prefer to give to students feedback…*", the learner erroneously used the preposition 'to', which was likely the result of the significant differences between Turkish and English syntax in the positioning and marking of objects as suffixes are frequently used in Turkish to distinguish between direct and indirect objects. This finding is consistent with that of Babanoğlu (2011), who examined the effect of L1 on the dative alteration structures produced by adult Turkish learners of English through various learner data derived from tests and corpus data. Research findings revealed that there was clear evidence of *L1 transfer* in terms of the use of dative structures.

Regarding the pattern G8(B) (*Verb + D O + to + I O*)*, L1 transfer* errors were highly prevalent, with 6 occurrences out of 8 (75%). However, in overall terms, this error type accounted for only 2.56% of all errors across different patterns. This implies that this issue seems to be pattern-specific rather than a generalised problem, as three-quarters of the errors within this pattern were caused by direct influence from the learners' native language (L1). This suggests that learners may have struggled to apply this structure due to L1 influence correctly. For instance, in the following sentence, "*It is essential for students to give attention to the lesson…*", the direct translation of Turkish expressions into English was the cause of this inaccuracy. The Turkish verb '*vermek*' (give) is used

in the comparable phrase '*derse dikkat vermek*' (literally: '*give attention to the lesson*'), which Turkish learners may mistranslate as '*give*' in English. This discrepancy arises since the verbs '*pay attention*' and '*give attention*' are used differently in Turkish and English to convey the same notion. This result contributes to the widespread impact of L1 interference in language acquisition, particularly with regard to syntactic and collocational elements (Lopez, 2016).

As for the pattern G8(N) (*verb + object + complement*), the results indicated that there was a high proportion of *L1 transfer* errors, with 15 instances out of 22 (68,18%), making it the most dominant type of error within this pattern. Overall, these errors accounted for 6,41% of all L1 transfer errors across the patterns. In spite of the fact that the rate as a whole remained quite low, this demonstrates that learners seemed to struggle to internalise the English-specific structure of *verb + object + complement* due to a direct influence from the learners' native language. This is consistent with the research by Swan (2005), Taouis and Pérez (2019) and Kim and Seo (2022), who highlighted that discrepancies in complement placement and selection stem from variations in complement structures between learners' L1 and the target language, potentially influencing obstacles in English language learning. For instance, in a sentence from the learner corpus, "*Teachers make easier understand by using different methods.*", the learner excluded the dummy pronoun '*it*' because their native language grammar lacks such syntactic features. Turkish syntax does not necessitate placeholders like '*it*', resulting in errors such as 'make easier understand' instead of '*make it easier to understand*'. This implies that students just took semantic meaning into account, failing to properly account for syntactic restrictions, which led to frequent mistakes in complement structures.

In pattern G8(O) (*Verb + Object 1 + Object 2*), the dominant type of error was *L1 transfer* errors with 4 out of 7 instances (57,14%). Overall, these mistakes made up 1,71% of all *L1 transfer* errors in various grammatical patterns. According to this, learners were more likely to make *L1 transfer* errors with other patterns; however, they had trouble with the proper syntactic structure of English double-object formulations, as seen by the fact that over half of the errors were caused by L1 impact. This finding is in line with He's (2022) study, which suggested that most of the erroneous English double object constructions produced by Chinese EFL learners resulted from transfer by L1. This shows that it may

be difficult for learners to tell the difference between using prepositions and placing objects directly or indirectly. In the sentence, "*In traditional classroom settings, education requires students attendance.*", the error likely indicates an instance of *L1 transfer* from Turkish, where possessive forms may be suggested or constructed differently than in English. In Turkish, the possessive link between '*students*' and '*attendance*' may not necessitate the explicit application of a possessive marker analogous to the English apostrophe + s ('s) or possessive adjective (*their*). This underscores the impact of Turkish syntax on learners' English proficiency, especially in multi-object constructions, where explicit grammatical links are necessary in English but may be implicit in Turkish.

Regarding the G3 (*noun + that clause*) pattern, the results revealed recurring problems, mostly with L1 interference. Errors in the *L1 transfer* category were observed in 7 cases out of 15 (46,67%) with an overall rate of 2,99%. Although the errors in this pattern did not represent a major proportion of L1 transfer errors across all structures, nearly half of them resulted from L1 influence, indicating that learners might have struggled with *noun + that-clause* constructions due to their native language structures. In the instance, "*The reason that because they leave school early is unknown.*", '*that because*' redundantly merges two phrases, possibly indicating L1 interference. Turkish frequently explains causes with conjunctions like 'çünkü' (because) or comparable constructs. Learners may translate these constructs straight into English, which could lead to redundant or improper wording, demonstrating how L1 (Turkish) grammatical patterns interfere (Karagöz, 2010). This result is consistent with Kırkgöz's (2010) study, which examined typical mistakes made by Turkish English language learners and found L1 interference in structures with '*that*'. Similarly, Hidayati and Suryani (2020) examined mistakes committed by Indonesian EFL students and discovered that noun clause omissions and misinformation were common problems. They ascribed these challenges to the influence of learners' native language and the structural intricacy of noun clauses. These results, however, only partially align with the findings of Nguyen and Tran's (2019) study, which examined noun clause usage errors among Van Hien University senior English majors. Their study found four primary categories of errors and attributed them to intralingual reasons, including the *overgeneralisation* of English grammar rules and interlingual interference from Vietnamese. Given that Turkish and Indonesian students may rely on

different L1 structures than Vietnamese learners, the discrepancy in results could be caused by differences in the learners' linguistic backgrounds. Furthermore, Nguyen and Tran's (2019) focus on intralingual issues raises the possibility that *overgeneralisation* might be more prevalent in situations when students encounter complex English structures or lack adequate contextual practice.

The findings showed that in pattern G8(Q) (*Verb + (Object) + Wh-Clause/Wh-Phrase*), *L1 transfer* errors were the most common error type, occurring in 8 out of 20 cases (40%). Overall, 3,42% of all L1 transfer mistakes across all patterns were caused by these problems. While the overall rate does not represent a large proportion of all L1-related errors across all patterns, L1 transfer is likely to be a major issue in this specific pattern. The 40% rate within the pattern indicates a consistent difficulty in applying English *wh-clause* structures, particularly in indirect speech and embedded questions. In the given sentence, "*Students ask them where can they stay…",* the learner misplaced '*they*' (indirect object) and produced a deviant structure. Since Turkish does not have a distinct word order change for direct and indirect enquiries, Turkish students may apply the direct question word order when building indirect *wh-clauses*. This finding aligns with the results reported by Bohran (2024), who emphasised that the primary source of error in learners' production of wh-phrases stemmed from L1 interference. The researcher observed that learners often transferred syntactic structures or word order patterns from their first language when forming wh-questions or embedded wh-clauses in English. This negative transfer may have led to deviations from target-like structures, thereby hindering the accurate acquisition and use of wh-phrases in the EFL context. The study underscored the persistent influence of the native language on second language development, particularly in syntactically complex structures such as wh-constructions.

For pattern G8(S/s) (*Verb + Complement (Adj. or N)*), *L1 transfer* was also the most frequent type of error, with 8 out of 27 instances (29,63%). Overall, these errors accounted for 3.42% of all *L1 transfer* errors across different grammatical patterns. While they did not constitute a significant portion of L1 transfer errors across all grammatical structures, these errors impacted this specific pattern. Nearly one-third of the errors in this pattern stemmed from L1 influence, suggesting that learners might have found structuring *verb + complement* sentences in English challenging. In the following case, *"People who work long hours without breaks are more likely to feel stressful." 'stressful'* was used

instead of '*stressed*'. The learner made a lexical error by directly translating the Turkish adjective '*stresli*' as '*stressful*', not realising that '*stressful*' denotes an external stimulus rather than an interior emotion. In Turkish, adjectives like '*stresli*' and verbs such as '*hissetmek*' (to feel) frequently appear in collocations where the differentiation between states and situations (e.g., '*stressed*' versus '*stressful*') is less pronounced. This results in ambiguity when employing the correct forms in English. The finding that negative transfer from L1 is a dominant issue in terms of the use of *verb + complement* is in line with the findings of Ayar (2020) who found that with regard to verb-related errors and lexical errors, negative transfer from L1 was the most dominant source of error in writings of 7 university students. In a similar vein, Çelik et. al (2023) investigated the English opinion paragraphs of 19 Turkish high school students and found that the most common errors were related to verbs. This finding suggests that these errors may be the result of syntactic differences between Turkish and English, which causes students to transfer Turkish verb formation rules into their English writing. For Turkish students of English as a foreign language, these findings collectively highlight the predominance of *L1 transfer* as a dominant cause of verb-related and lexical errors for these students. These difficulties are likely to be worsened by the structural differences that exist between Turkish and English, particularly in the conjugation of verbs and the utilisation of vocabulary.

As for pattern G8(P) *(Verb + (Object) + Adverbial) L1 transfer* errors were the second most prominent error, occurring in 14 cases out of 56, with a percentage of 25%. In overall terms, the rate was 5,98%. This demonstrates that *L1 transfer* could be a major concern in this pattern despite the fact that its overall rate suggests learners may encounter more substantial L1-related challenges in other grammatical structures. This result also indicated that learners encountered difficulty in arranging *verb + object + adverbial* sentences in English, as L1 interference was responsible for one out of every four errors. For example, in the following sentence, "*Children can remember the moments beautifully ...*", '*beautifully*' was employed where '*fondly*' was more suitable for characterising memory recollection in English. In Turkish, adverbs such as '*güzelce*' (*beautifully*) have expansive, adaptable meanings and are employed in situations where certain adverbs like '*fondly*' would be favoured in English. This finding is in line with that of White (1991) and Gledhill (2005), who emphasised that learners often misuse adverbs in sentences and

that the source of learners' errors was *L1 transfer* regarding the use of this structure. Learners directly translate from their native language, resulting in unnatural or improper adverb placement or selection in English.

Regarding certain patterns, the total number of *L1 transfer* errors was relatively lower. To illustrate, in G8(H) (*Verb + Object + to Infinitive*), there were 3 instances out of 22 (13,64%). Overall, these errors accounted for just 1,28% of all L1 transfer errors across different grammatical patterns, highlighting the fact that *L1 transfer* in G8(H) was not prevalent across all patterns and within this specific pattern. The instances of *L1 transfer* errors in the learner corpus indicated that L1 syntax may have occasionally affected their decisions. In the sentence, "*The new program provided students to take advanced courses*", the corresponding verbs for '*provide*' (sağlamak) and '*enable*' (imkan tanımak) frequently overlap in meaning and usage. In Turkish, verbs such as '*sağlamak*' can express both '*to provide*' and '*to enable*', resulting in ambiguity in English; therefore, they may erroneously translate the structure by appending '*to take*' to verbs such as 'provide', resulting in ungrammatical constructions in English. There is a similar case for pattern G8(R) (*It+ Verb + Object + to Infinitive/that-clause*). Concerning this pattern, *L1 transfer* errors occurred in 2 cases out of 12 (16,67%), with an overall rate of 0,08%. In the following case, "*The new online platform is beneficial because it provides us to receive training from experts around the world.",* the learner used the word '*provide*' instead of '*allow*'. This error occurred since '*provide*' in English does not accept the structure '*to receive*' without an indirect object. However, in Turkish, verbs like '*sağlamak*' (provide) are often used in structures that may not require an indirect object or infinitive distinction, leading learners to misuse '*provide*' in English by mapping it directly from Turkish. Consequently, *L1 transfer* errors in the learner corpus suggested that learners' decisions were intermittently influenced by L1 structure despite their reliance on their L2 proficiency.

There were no instances of *L1 transfer* errors in a total of 12 patterns: G2, G6, G7, G8(C), G8(E), G8(F), G8(G), G8(I), G8(J), G8(K), G8(L), and G8(M). The absence of *L1 transfer* errors indicates that learners seemed not to depend on their first language (L1) to establish these structures, which implies either a more profound understanding of the target language's norms or a reduced probability of negative L1 influence in these specific patterns. This could be attributed to factors such as explicit instruction, frequent exposure

to correct usage, or structural similarities between the learners' L1 and the target language in these cases. The absence of *L1 transfer* errors in these patterns may also suggest that learners have internalised the correct grammatical forms, thereby reducing the necessity for L1-based processing (Al-Sobhi, 2009). Learners progressively internalise the correct forms when they encounter consistent grammatical structures on a regular basis. Over time, this eliminates the necessity of cognitively translating prior to generating language, thereby facilitating the spontaneous use of a second language.

5.2.3.2 False Concept Hypothesised Errors

Overall, the total number of *false concept hypothesised* errors across all patterns showed that G8(P) *(Verb + (Object) + Adverbial)* had the highest percentage of this type of error, at 24,78% (28 occurrences out of 113) and the second-highest score (10,62%; 12 out 113) was observed in G1 (*noun + preposition*). On the other hand, the lowest percentages were seen in patterns G8(K) (*Verb + Possessive + V-ing*) (1 out of 113; 0,88%) and G2, G4, G6, G8(G), G8(J), G8(M) (2 out of 113, 0,88% each) (See Table 4.8, p. 71). This distribution confirms that *false concept hypothesised* errors were not uniformly distributed across all structures and were a concern in certain areas. However, they did not seem to be equally problematic in all grammatical patterns. This implies that certain patterns may necessitate more conceptual clarification than others.

When the prevalence of the *false concept hypothesised* error was analysed within each pattern, the results revealed that it was the most dominant source of error within specific patterns: G8(H) *(Verb + Object + to Infinitive)*, G8(L) (*Verb + that-clause*), and G8(P) *(Verb + (Object) + Adverbial)* (See Table 4.10, p. 74). Among these, *false concept hypothesised* error was mostly observed in pattern G8(P) *(Verb + (Object) + Adverbial)*, with 28 occurrences (out of 113; 24,78%) and was the most dominant type of error within this pattern (out of 56, 50%), indicating the dominant impact of it in both overall and specific terms. The results showed that learners likely exhibited a notable misunderstanding of how verbs interact with adverbials, making this the most problematic area in their language acquisition process. To illustrate, in the sentence, "*The advantages of a new system cannot be achieved <u>complete</u> without proper planning,*" the term '*complete*' was erroneously employed as an adverb, likely stemming from a misunderstanding of the English convention of forming adverbs by appending '*-ly*' to

adjectives. This can be attributed to the observation made by Lightbown and Spada (2013), who emphasised that learners frequently have difficulties in differentiating between adjectives and adverbs in a second language due to insufficient exposure to exceptions and irregularities. This implies that learners may have encountered significant challenges in comprehending the interaction between verbs and their adverbial modifiers accurately. In particular, these errors were likely the result of learners' incorrect hypotheses about the interaction between verbs and adverbials, which resulted in misinterpretations or improper constructions.

The results revealed that for pattern G8(H) *(Verb + Object + to Infinitive) false concept hypothesised* occurred in 9 cases and was the most frequent type of error within this pattern (40,91%) and had an overall rate of 7,96%. This reveals that it was a major issue in the acquisition of this pattern, indicating that learners frequently misapplied or misunderstood the syntactic rules governing verb complementation in English. These errors frequently arose when learners needed to differentiate between verbs that require gerunds and those that necessitate the use of '*to + infinitive*'. For instance, in the given sentence, "*The new law permits agencies <u>to listening</u> to private conversation*s," the learner mistakenly used a gerund after the verb '*permit*' which necessitates '*to + infinitive*' as a result of a misinterpretation of the *Verb + Object + to Infinitive* structure. The frequency of these errors within this pattern suggests that learners may have developed incorrect assumptions about verb complementation rules. Additionally, as Zinteng et. al (2023) stated, learners' misinterpretations of verb complementation often stem from the misapplication of rules, leading to incorrect structures involving objects and infinitives in English.

In the utilisation of pattern G8(L) (*Verb + that-clause*), there was a prevalent occurrence of *false concept hypothesised* errors, constituting 7 of the 14 errors (50%), which signifies that learners often misapplied or overextended their comprehension of *that-clauses*. The 6.19% overall error rate means that while not the most dominant issue, *false concept hypothesised* errors in that-clause constructions seemed to be a recurring problem in learner writing. Within the G8(L) pattern specifically, half of all errors were *false concept hypothesised* errors, meaning that a considerable number of learners struggled with the proper application of that-clauses. These errors occurred when learners made deviant assumptions about *verb + that clause* construction. In the sample case, "*A lack of practice*

often leads that students do not to acquire the necessary skills…", the learner erroneously assumed that '*that*' is invariably required after specific verbs, resulting in repetition or ungrammatical constructions, probably arising from their syntactic complexity. These findings suggest that learners are likely to frequently misapply or overextend the rules governing *verb + that-clauses*, leading to syntactic errors.

In some cases, although the total number of *false concept hypothesised* errors was lower in overall terms, this source of error remained the most prevalent source of error within patterns G8(C) *(Verb + D O +for + I O =Verb + I O + D O),* G8(J) *(Verb + Object + V-ing),* G8(M) (*Verb + object + to be + complement*)) and G8(K) (*Verb + Possessive + V-ing*). This indicates that regarding these patterns, the overall impact of the *false concept hypothesis* was low, suggesting that this pattern seemed not to be a widespread issue for learners. These errors were the primary source of errors, indicating a conceptual misunderstanding specific to these patterns. To illustrate, the principal obstacle regarding the pattern G8(C) *(Verb + D O +for + I O =Verb + I O + D O)* was the *false concept hypothesised* with 3 out of 4 (75%) within the pattern and 3 instances out of 113 (2,65%) overall, meaning that most learners who struggled with this structure might have done so because of incorrect assumptions about *verb-object-preposition* compositions. In the sentence, "*Some believe society should show more respect for famous people,*" '*respect*' was erroneously coupled with '*for*' rather than '*to*'. These errors indicate that the learner seemed to be confused about the appropriate preposition needed in specific *verb-object-preposition* formulations. This error can be explained by the fact that collocations provide significant challenges for non-native speakers because of their arbitrary characteristics (Nesselhauf, 2005). This finding is in line with that of Taouis and Pérez (2019), who examined noun and direct object collocations in the written English of Spanish learners, uncovering faults in the application of verb patterns with indirect objects. They highlighted that students often encounter difficulties assigning suitable responsibilities to objects, resulting in deviations from natural standards. Nevertheless, the results of this research are in direct opposition to those of Millaku (2015), who investigated the utilisation of direct and indirect objects in Albanian and English. The researcher emphasised that learners frequently depend on L1 transfer in constructions such as *Verb + Direct Object + for + Indirect Object*. This disparity may be attributed to the structural

distinctions between the learners' native language (L1) and English (L2), which affect the formation and comprehension of such patterns.

Regarding G8(J) *(Verb + Object + V-ing)* pattern, the total number of errors was low, with only 2 instances (1,77%), indicating in overall terms, *false concept hypothesised* error was not prevalent and contributed very little to the total number of these errors. However, within G8(J), it was the dominant error type (2 out of 3 errors, 66,67%), implying that learners' difficulties with G8(J) were predominantly the result of conceptual misunderstandings rather than other types of errors. This may indicate learners' challenges in employing the appropriate possessive form and gerund usage inside *verb-object-gerund* formations. In the learner corpus, *false concept hypothesised* errors arose when learners misinterpreted the possessive link necessary between the object and the gerund. For example, the learner employed '*them thinking*' rather than the grammatically accurate '*their thinking*' in this sentence, "*By dematerialising them thinking, individuals focus more on ideas, experiences, and values*". In this case, the learner may have conflated the possessive form needed before a gerund with the object pronoun generally employed in other verb-object structures due to a misinterpretation of possessive constructions in the English language. As Swan and Smith (2001) stated, learners' difficulty in recognising the relationship between non-finite verb forms and possessive determiners frequently results in errors involving gerunds, particularly in languages where such distinctions are less emphasised.

There is a similar case for pattern G8(M) (*Verb + object + to be + complement*), there were 2 out of 3 occurrences of *false concept hypothesised* errors (66.67%). While *false concept hypothesised* errors in the G8(M) pattern were not frequent in absolute numbers (1,77%), their high rate within the pattern suggests that learners who attempted to use this structure might have misinterpreted its rules. The primary difficulties arose from misunderstanding how '*to be*' functions. In the given instance, "*Parents want their children to be succeed at school,"* the student erroneously employed the base verb '*succeed*' in place of the adjective '*successful*', demonstrating a misunderstanding of the role of '*to be*' as a copula that connects the subject to a descriptive complement. Learners may have erroneously assumed that '*to be*' can introduce a verb in all circumstances, leading to the misconception of its grammatical role, which may result from a limited understanding of the target language system, occurring when learners are tasked with

performing beyond their current linguistic proficiency. However, this finding is not consistent with that of Yip (1995), who emphasised that errors in structures such as G8(M) are frequently affected by the learner's L1, Chinese. This inconsistency may be due to the grammatical and structural differences between Turkish and Chinese, which highlights the significance of evaluating the typological closeness of L1 and L2 in comprehending the influence of L1 transfer on language acquisition.

Within pattern G8(K) (*Verb + Possessive + V-ing*), there was only 1 error and it fell under the category of *false concept hypothesised*. In overall terms, the rate was exceedingly low, 0,88%, indicating that when learners struggled with G8 (J), it was primarily owing to conceptual misunderstandings rather than other types of errors. *False concept hypothesised* error occurred in the sentence "*Teachers dislike their students arrive late to the class.*", indicating a misinterpretation, not realising that the gerund '*arriving*' is needed as the object of '*dislike*'; the student could hypothesise that the possessive noun '*their students"* must be followed by the base form of the verb '*arrive*'. This can be explained by Lightbown and Spada's (2013) assertion that these errors occur when learners form erroneous assumptions about the target language, frequently shaped by inadequate or overly simplistic instruction. As mentioned earlier, learners infrequently engage with this structure or prioritise simpler alternatives. This corresponds with Bardovi-Harlig (2002), who claimed that students usually choose simpler, more common patterns over rare, complex forms.

In some cases, *false concept hypothesised* were also frequent as they were observed as the second most dominant type of error within certain patterns (G3 *(noun + that clause)* and G8(O) *(Verb + Object 1 + Object 2)).* To illustrate, the results revealed that in pattern G3 *(noun + that clause), false concept hypothesised* errors were the second most dominant type of error with 5 instances out of 12 (33,33%). In overall terms, the error rate was 4,42%, suggesting that it was not among the most frequent in overall language acquisition but was still present. In the following sentence: "*The possibility that if online education will become permanent…*" the functions of '*that*' and '*if*' led to erroneous clause construction. This misconception indicated that learners may have mistakenly assumed that '*that*' and '*if*' can be used interchangeably or coupled within English *noun + that-clause* constructions. This error can be attributed to the fact that students may add unnecessary conditional clauses to a statement and make it overly complicated (Swan,

2005). This mistake also aligns with what Richards and Schmidt (2013) describe as a case in which learners have a misunderstanding of the regulations regarding clause construction and conditionality. Princess and Syraiff (2021) observed that misunderstandings regarding the use of *noun + that-clause* constructions in academic and formal writing can result in substantial difficulties in constructing sentences that are both clear and precise. The learners' capacity to communicate effectively in formal writing contexts may be impeded by their failure to employ these structures correctly.

As for the pattern G8(O) *(Verb + Object 1 + Object 2), false concept hypothesised* errors had a percentage of 42,86% (3 out of 4 instances), and the overall rate was relatively lower: 2,65%. This indicates that this type of error in this pattern was high, and learners seemed to struggle with understanding how to correctly structure double-object verbs, leading to frequent misinterpretations of verb-object-object constructions. For example, in the following sentence, "*Experts give to them the necessary information…*", the learner developed a flawed comprehension of the expression of possession and noun modification in English, probably assuming that all indirect objects in English must be introduced by a preposition. This can be attributed to the fact that learners may form inaccurate hypotheses regarding L2 structures as they develop their interlanguage, a transitional linguistic system that is influenced by both L1 and L2. This process has the potential to result in errors in the formulation of double-object verbs (Tarone, 2006).

Regarding pattern G1 *(noun + preposition)*, *false concept hypothesised* errors occurred in 12 cases out of 65 (18,46%), and the overall proportion was 10,62%. This suggests that while false concept hypothesised errors in noun–preposition structures were not the most dominant error type overall, they still occurred with sufficient frequency to be considered a notable learning challenge. Specifically, within the G1 pattern, 18,75% of the errors stemmed from incorrect assumptions about noun–preposition pairings, indicating that a considerable number of learners may have misinterpreted the syntactic or semantic functions of prepositions following nouns. In this instance, "*Students can concentration on problems more…",* the learner confused the verb '*concentrate*' with the noun '*concentration*', presuming that verbs such as '*concentrate*' can function as nouns without morphological modifications. This finding is in line with that of Visetti and Cadiot (2002) which examined prepositions' semantic instability and demonstrated how their polysemy frequently results in misunderstandings in *noun + preposition* formations.

It supports the notion that learners' conceptual misunderstandings lead to mistakes. This implies that learners may have encountered difficulties in accurately associating nouns with their corresponding prepositions, resulting in errors in word selection and usage. These errors are likely the result of a failure to properly apply prepositional rules or a misinterpretation of word forms.

This type of error was also notable in pattern *G8(N) (Verb + Object + Complement),* with 4 instances out of 22 (18,18%) indicating the misinterpretation of the distinctive structure of the G8(N) pattern. Overall, its rate was 3,54%, suggesting that while *false concept hypothesised* errors in this pattern were not highly frequent overall, they were notable within this specific pattern, suggesting that learners seemed to recognise this structure but struggle with its proper application. In the given sample, "*Society considers important them because their contributions positively impact the community.*", the learner erroneously positioned the compliment '*important*' before the object '*them*', potentially due to a misapprehension of English syntactic structure. Insufficient exposure to genuine applications of *verb + object + complement* structures may have resulted in inadequate internalisation of the proper word order. As Domede (2024) stated, grammatical mistakes occur frequently when using complicated structures and have been shown to impair students' academic performance by making their assignments harder to read and less coherent.

In certain cases, both the overall rate and the proportion of these errors were lower (G2 *(noun + to + infinitive),* G4 *(preposition + noun),* G5 *(adjective + preposition),* G6 *(adjective + to infinitive)*, G8(D)/(d) *(Verb + (Object) + Preposition + Object),* G8(G) *(Verb + V-ing),* G8(E) *(Verb + to + Infinitive)*. This indicates that these patterns had a lower contribution to the overall or specific problem of *false concept hypothesised* errors. Among these patterns, the highest proportion was observed in pattern G4 *(preposition + noun)*, with 2 instances out of 45, accounting for only 4,44% of errors, with an overall rate of 1,77%. This means that most errors in this structure might have arisen from other sources (e.g., lexical confusion or word choice errors rather than a fundamental misunderstanding of grammatical rules). This indicates that learners possess a relatively precise comprehension of *preposition + noun* structures. In the instance, "*The population continues to increase despite of the fact that many young people are moving abroad…*", the learner erroneously assumed that '*of*' is necessary after '*despite*', signifying a

misapprehension of the role of '*despite*' as a preposition. Students may have encountered '*in spite of*' and erroneously generalised the structure to '*despite*'. This error may occur when learners are uncertain about the collocational rules of English and add superfluous prepositions, such as '*of*'. On the other hand, G8(G) *(Verb + V-ing)* had the lowest frequency of *false concept hypothesised* error with 2 occurrences out of 8 (16,67%) within the pattern. The overall rate of this type of error was 1,77%. Given the low incidence of these errors in G8(G), it appears that learners did not seriously misinterpret the requirement that the -ing form comes after some verbs. Nevertheless, the two instances in the learner corpus may underscore learners' difficulty in discerning whether verbs require gerunds as opposed to infinitives, indicating insufficient internalisation of these principles. In the instance, "*Teachers often need to guiding students about complex topics…*", '*need*' necessitates an infinitive instead of a gerund, signifying a misunderstanding of the verb's argument construction. This error might result from a misinterpretation of the appropriate use of the infinitive (to guide) compared to *V-ing*. As suggested by Lightbown and Spada (2013), learners' errors are frequently the result of their incorrect interpretation of target language input, particularly when coping with intricate grammatical categories such as infinitives and gerunds.

There were no instances of *false concept hypothesised* error in patterns G7 (Adjective + that-clause), G8(A) (Verb + D O + to + I O = V + I O + D O), G8(B) (Verb + D O + to + I O) and G8(R) (*It + Verb + Object + to Infinitive/that-clause*). The lack of *false concept hypothesised* errors in these structures may suggest that learners did not have significant conceptual misconceptions about the grammatical requirements of the structure. These errors arise when learners develop inaccurate hypotheses about the target language as a result of a defective understanding of the language's distinctions. The absence of such errors indicates that learners seem not to struggle with conceptual misconceptions regarding their formation and usage, which may be attributed to adequate exposure or effective instruction (Khansir, 2012).

5.2.3.3 Overgeneralisation errors

In overall terms, across all patterns G1-G(S/s), the total number of *overgeneralisation* errors indicated that this source of error was mostly observed in G8(F) *(Verb + bare infinitive)* with a percentage of 19,35% (12 cases out of 62). G8(G) *(Verb + V-ing)* was

the second highest (12,90%; 8 out 62) across all patterns, suggesting that learners tended to apply a single rule too broadly, incorrectly extending certain grammatical structures. In contrast, the lowest percentages were seen in patterns G8(J) *(Verb + Object + V-ing)* and G8(M) (*Verb + Object + to be + Complement*) (2 out of 62, 1,61%) (See Table 4.8, p. 71). This shows that *overgeneralisation* errors were distributed more frequently among other patterns**,** confirming that the distribution of these errors was not consistent across all grammatical constructions.

Regarding the prevalence of *overgeneralisation* errors, specifically within the patterns, the results revealed that the highest percentage of these errors was seen in G6 *(adjective + to infinitive),* G8(F) (*Verb + bare infinitive*)*,* G8(G) *(Verb + V-ing)*, G8(I) *(Verb + Object + Infinitive)* and G8(R), (*It+ Verb + Object + to Infinitive/that-clause*) (See Table 4.10, p. 74). To illustrate, this source of error was observed in pattern *G6 (adjective + to infinitive)* in 5 out of 7 instances (71,43%). Overall, the proportion was 8,06%, indicating that *overgeneralisation* errors in the G6 pattern were relatively moderate overall, but extremely prevalent within this specific pattern. This suggests that learners seemed to misapply grammatical rules when using the *adjective + to-infinitive* structure, most likely due to confusion with similar adjective or infinitive constructions. For instance, in the following sentence, "*New vocabulary is often easier <u>to learning</u> when used in real-life contexts.*", the learner applied a rule beyond its relevant context and erroneously employed the gerund '*learning*' instead of the appropriate infinitive '*to learn*'. This may have occurred due to the prevalent use of gerunds in different situations. Students may link verbs such as '*learning*' with habitual application, failing to acknowledge that the infinitive form is required after '*to*'. Due to confusion with other structures, such as verb + gerund patterns (e.g., *look forward to learning*), learners may have incorrectly used a gerund form following 'to' in this instance. When students apply rules from one context (*verb + gerund*) to another (*adjective + to + infinitive*), they might engage in *overgeneralisation*, according to Ellis (2003).

Within G8 (G) *(Verb + V-ing)* pattern, 8 out of 12 errors (66.67%) resulted from *overgeneralisation*, making it the most common type of mistake in this structure. In overall terms, 8 out of all errors (12.19%) were *overgeneralisation* errors related to this structure, indicating that learners partially acquired the gerund form but did not acknowledge exceptions, as in the following sentence, "*After careful consideration, he*

decides sending additional resources…". In this case, the learner erroneously used the gerund form '*sending*' instead of '*to send*'. This aligns with the explanation made by Ellis (2003), who asserted that such errors occur when learners apply a learned rule to contexts in which it is not applicable, such as using *V-ing* with verbs like '*decide*'. In general terms, the high prevalence of *overgeneralisation* errors in the G8(G) pattern is consistent with that of Unggul and Gulö (2017), who investigated *gerund* and *to-infinitive* usages made by students in their argumentative essays and concluded that the most common problem in using *gerund* and *to-infinitive* was misformation leading to *overgeneralisation*, a type of error closely related to intralingual errors.

Regarding the pattern G8(F) (*Verb + bare infinitive*) *overgeneralisation* was also the most prevalent error with 12 instances (48%). In overall terms, the rate of it was 19,35%, suggesting that *overgeneralisation* errors in the G8(F) pattern were prevalent overall and even more dominant within this specific pattern. This shows that learners seemed to frequently apply rules to situations where they were inapplicable, likely due to incorrect generalisations from other verb patterns. In the context, "*Building good habits will helps them succeed…*", the learner erroneously inserted a '*-s*' to '*helps*' incorrectly applying the rule of subject-verb agreement from earlier instances. In this case, the learner might have overgeneralised subject-verb agreement principles and incorrectly applied third-person singular -s to verbs that follow modals such as '*will*'. This is consistent with the assertion made by Brown (2000) indicating that *overgeneralisation* is a prevalent intralingual error, characterised by the propensity of learners to excessively apply rules they have partially acquired.

As for the pattern G8(R), (*It+ Verb + Object + to Infinitive/that-clause*), *overgeneralisation* errors were one of the most dominant errors with 5 out of 12 instances (41,67%) and the overall proportion of this type of error was 8,06%, highlighting that *overgeneralisation* errors in this pattern were not highly frequent overall; however, they were frequent within this specific pattern. This suggests that learners frequently used this structure erroneously, most likely due to confusion between infinitive and that-clause constructions or other verb complementation patterns. For instance, in the given sentence, "*Meditation is a powerful practice because it helps us relaxing.*", the learner incorrectly applied the gerund (-ing form) in the situation requiring an infinitive. Students may link all verbs after '*help*' with the gerund, likely due to insufficient exposure to accurate

usages. Learners may generalise the gerund structure seen in other contexts, such as '*enjoy relaxing*', to verbs like '*help*', which do not permit it. This finding aligns with that of Chen (2024) who analysed error rates in syntactic constructions like G8(R). It was reported that these errors often stemmed from *overgeneralisation* and insufficient understanding of English clause structures. Nevertheless, this is inconsistent with the findings of Taouis and Pérez's (2019) study, which attributed these errors to L1 interference and inadequate instructional focus on complement clauses. As mentioned above, the focus of Taouis and Pérez's (2019) study was on gerund and infinitive errors in Spanish learners of English; therefore, this discrepancy can also be attributed to the fact that Spanish and Turkish differ in terms of their grammar and structure.

Concerning the pattern G8(I) *(Verb + Object + Infinitive), overgeneralisation* was the most dominant type of error. It was observed in 7 out of 18 instances within the pattern, constituting 38,89% of this type of error, and in overall terms, the rate was 11,29%. The high prevalence of *overgeneralisation* errors in the G8(I) pattern might suggest that learners struggled to distinguish between verb patterns requiring a to-infinitive and those not. In the sentence, "*They planned a hands-on activity to <u>let them to discover</u> new concepts on their own."* the learner unnecessarily attached the infinitive marker '*to*' after the verb '*let*', which did not necessitate it. This is a flawed use of the principle that '*to*' must precede all infinitives. This can be ascribed to Lightbown and Spada (2013), who emphasised that high-frequency norms, such as '*to*', are particularly susceptible to *overgeneralisation* in learners' language.

In some cases, the rate of *overgeneralisation* was relatively high, whereas it was lower in overall terms (e.g. G8(J) *(Verb + Object + V-ing) and* G8(M) (Verb + Object + to be + Complement). Regarding G8(J) *(Verb + Object + V-ing)*, it was seen in 1 instance out of three (33,33%) within the pattern, and the overall rate was 1,61%, highlighting learners' difficulties in employing the appropriate possessive form and gerund usage inside *verb-object-gerund* formations. *Overgeneralisation* occurred in the instance, "*Professionals who frequently keep their clients <u>wait</u> risk their reputation and…*", where the learner erroneously simplified the structure by employing the basic verb form '*wait*' rather than the gerund '*waiting*'. This may indicate a misapprehension of the grammatical principle that specific verbs, such as '*keep*', necessitate the gerund form following the object. In this case, learners may have correlated '*keep*' with base verb forms because of their

exposure to more uncomplicated phrases or constructs where the base form is suitable (e.g., *keep going*).

There is a similar case for pattern G8(M) (*Verb + Object + to be + Complement*). An *overgeneralisation* error was observed in only one instance out of three total errors within the G8(M) pattern, accounting for 33.33% of errors in that specific category and representing an overall error rate of 1.61%. This suggests that while *overgeneralisation* was relatively rare across the dataset, it constituted a notable proportion of errors within this particular pattern. Such errors likely stem from the misapplication of grammatical rules, particularly in distinguishing when the verb 'to be' is required in verb–object–complement constructions. An illustrative example of this is the sentence: "*Experts suggest it to be done with caution for safety,*" where the structure deviates from standard usage due to overgeneralised rule application. This error illustrates the *overgeneralisation* of the '*to be*' structure when the learner erroneously applied it to circumstances dictated by verbs necessitating *bare infinitives*. In this instance, '*suggest*' does not necessitate '*to be*'; however, the learner might have mistakenly applied the rule for causative verbs (e.g., *want it to be done*) to verbs such as '*suggest*'. As Bourke (2007) stated, the challenges that learners encounter as a result of the subtle distinctions in verb patterns and the limited effectiveness of formal norms in teaching these structures are underscored by verbal complementation errors.

As for pattern G8(H) (*Verb + Object + to Infinitive*), in 5 out of 22 instances (22,73%) *overgeneralisation* was observed within this pattern, making it a notable but not dominant issue within this structure. 5 out of all errors (8.06%) were *overgeneralisation* errors related to the G8(H) pattern, indicating that this structure was a moderate source of mistakes in overall learner performance. This suggests that when utilising the *Verb + Object + to-Infinitive* structure, learners seemed to make grammatical errors frequently. This is probably because they were confused by similar verb complementation patterns or had the wrong idea about when 'to' should come before an infinitive. Learners applied the *to + infinitive* structure to verbs such as '*prevent*', which necessitate prepositions followed by gerunds ('*from living*') as in the given sentence: "*Strict regulations can sometimes prevent people to live in certain areas*". Learners may not have recognised that verbs such as '*prevent*' require prepositional phrases and gerunds rather than *to + infinitive*. This type of error is consistent with Nesselhauf's (2005) explanation,

highlighting that such errors frequently arise from inadequate familiarity with verb-specific collocational rules in authentic language usage.

Regarding patterns G8(S/s) *(Verb + Complement (Adj. or N))* and G*8(E) (Verb + to + Infinitive), overgeneralisation* errors were moderately frequent within this specific structure, yet relatively low overall. 4 out of all errors (6.45%) were this source of error related to the G8(S/s) *(Verb + Complement (Adj. or N))* pattern, meaning that these errors are not highly frequent across all structures but still occur. However, within the G8(S/s) pattern itself, 14.81% of errors were due to *overgeneralisation*, showing that learners tended to misapply rules in a noticeable number of cases within this structure. In the given example, "*It's important to stay calm when another person feel panicked,…*" the learner failed to acknowledge the specific requirement of third-person singular agreement in English, resulting in an *overgeneralisation* of base verb forms. This suggests that learners occasionally misapplied grammatical rules when forming verb-complement constructions, likely due to confusion between different verb complementation patterns, such as using the base verb form for all subjects. For pattern G8(E) *(Verb + to + Infinitive) overgeneralisation* errors were observed in 6 out of 45 cases within the pattern (13,33%) and the overall rate was 9,68%. This suggests that learners seemed to misapply grammatical rules when using this structure, arising from the erroneous application of a rule across several circumstances. In the given sentence, "*Experts suggest to study on a particular subject.",* the learner improperly extended the use of *verb + to + infinitive* rule to the verb '*suggest*' instead of a gerund form. This finding is consistent with the observation made by Kim and Yoo (2015), who argued that many learner errors stem from insufficient knowledge of subcategorisation, the syntactic information that governs how verbs and other lexical items select their complements. In particular, learners often struggle to recognise which structures (e.g., infinitives, gerunds, clauses, or prepositional phrases) can appropriately follow specific verbs, resulting in the incorrect formation of collocational patterns. Such errors highlight a gap in learners' grammatical competence, specifically in their ability to internalise and apply the syntactic constraints associated with particular lexical items.

A total of 14 patterns did not exhibit *overgeneralisation* errors: G1, G3, G4, G5, G7, G8(A), G8(B), G8(C), G8(D/d), G8(K), G8(N), G8(O), G8(P) and G8(Q). Learners' lack of *overgeneralisation* errors implies that they may exhibit a more precise understanding

of the application of these structures by refraining from applying broad or incorrect general principals. Additionally, the absence of *overgeneralisation* errors may imply that these structures are incompatible with rule extension, which would suggest that learners did not perceive them as adhering to a predictable pattern that could be falsely generalised (Lightbown and Spada, 2013). This discovery emphasises the possibility that these grammatical forms are less susceptible to cognitive processing errors associated with *overgeneralisation*.

5.3.3.4 Incomplete application of rules errors

Across all patterns, the total number of *incomplete application of rules* errors indicated that this source of error was mostly observed in G8(D)/(d) *(Verb + (Object) + Preposition + Object)* with a percentage of 32,46% (37 cases out of 114). G8(E) *(Verb + to + Infinitive)* was the second highest (20,18%; 23 out 114). On the other hand, the lowest percentages were seen in patterns G8(B) *(Verb + D O + to + I O)* and G8(C) *(Verb + D O +for + I O =Verb + I O + D O)* (1 out of 114, 0,88%) (See Table 4.8, p. 71). This indicates that, among other patterns, errors resulting from incomplete rule application were dispersed more frequently.

Regarding the prevalence of *incomplete application of rules* errors specifically within the patterns, the results revealed that the highest percentage of errors was seen in patterns G2 *(noun + to + infinitive)*, G7 (*adjective + that clause*), G8(E) *(Verb + to + Infinitive)* (See Table 4.10, p. 74). The highest percentage was observed in G7 (*adjective + that clause*), with 6 out of 7 instances (85,71%). Overall, the error rate for this pattern was 5.26%, suggesting a partial understanding among learners of how and when to correctly construct sentences involving adjectives followed by a that-clause. Learners appeared to experience difficulty in forming the dependent clause, often omitting essential syntactic components, such as the subject pronoun. For instance, in the sentence *"It is certain that will cause problems,"* the omission of the pronoun *'it'* compromises the grammatical integrity of the clause. This pattern of omission may reflect inadequate syntactic processing, wherein learners recognise the need for a that-clause but fail to fully realise its structure. This finding aligns with previous research indicating that learners' inability to employ appropriate collocational patterns often leads to systematic grammatical errors. For example, Hashemi et al. (2012) conducted an investigation into collocational errors in the writing of university students and discovered that a significant contributor to these errors

was a lack of awareness of rule restrictions. Similarly, this is consistent with the research conducted by Aye (2020), which showed that learners' deficiencies in the appropriate application of this collocational pattern were the source of their errors. These studies indicate that language learners frequently encounter errors due to their inadequate comprehension and implementation of collocational patterns.

Regarding pattern G8(E) *(Verb + to + Infinitive)*, the results indicated that the most prevalent error type was *incomplete application of rules* (51,11%); the overall frequency of *overgeneralisation* errors was 23 out of 114 (20,18%). The high prevalence of *incomplete application of rules* in the G8(E) pattern may indicate learners' difficulties in completely implementing the syntactic rules of the G8(E) pattern, such as neglecting essential auxiliary or linking components (e.g., *to be*). In the given sentence, "*They don't need to anxious about the outcome, as everything is under control.*", the learner neglected the verb '*to be*', which might indicate a limited comprehension of the rule, coupled with an inability to implement it completely.

In pattern G2 *(noun + to + infinitive)*, there were 6 out of 12 instances of *incomplete application of rules* (50%), and the overall rate was 5,26%, suggesting that while its overall frequency across all errors was 5.26%, incomplete application of rules was a frequent issue in the G2 pattern. This indicates that learners might have grasped aspects of grammatical structure but did not fully implement it, neglecting essential components such as '*to*'. In the example, "*Effort help struggling students may include…*", the learner omitted '*to*', resulting in an ungrammatical construction. Students may have identified the *noun + to + infinitive* combination but neglected to insert '*to*' in the intervening position, particularly within complex sentence constructions. According to Brown (2000), such omissions may arise from restricted processing capacity or insufficient command of multi-step grammatical rules. Therefore, this absence indicates a deficiency in comprehending the proper structure of English *noun + to + infinitive* sentences.

In certain cases, this type of error was still prevalent as they were observed as the second most frequent one within some patterns (G8(I) *(Verb + Object + Infinitive)* and G8*(Q) (Verb + (Object) + wh- clause/ wh-phrase)*. For instance, in pattern G8(I) *(Verb + Object + Infinitive),* errors of *incomplete applications* arose in 6 cases out of 18 (33,33%), making it a prevalent source of error. While the overall rate was relatively lower (5,26%),

the results showed that learners seemed to frequently fail to fully apply grammatical rules required for *verb + object + infinitive* constructions. These errors involved incorrect word order, confusion between infinitive and gerund structures or when learners neglected essential components of the structure, such as the object pronoun '*them*' in the following sentence, "*Various methods let improve their talents in…".* This error resulted from omitting the object (*them*) after verbs such as '*let*', which led to the incomplete construction. This may signify a limited comprehension of the pattern and an incapacity to implement it entirely. In this case, learners may have concentrated on the verb and infinitive, neglecting the requirement for a clear object. For G8(Q) *(Verb + (Object) + wh- clause/ wh-phrase), incomplete application of rules* errors were also the second most prevalent type of error with 5 instances out of 20 (25%), while its overall frequency across all errors was 4.39%. The moderate frequency of incomplete application of rules in the G8(Q) pattern suggests learners might have acknowledged the necessity of a *wh-word* but did not adequately finalise the phrase. The absence of '*where*' in the example, "*The software organises videos by predicting they belong*", renders the sentence structurally and semantically deficient. Learners may have found it challenging to remember particular *wh-words* due to shortcomings in vocabulary or inadequate practice. As Doughty and Williams (1998) proposed, these errors may stem from insufficient procedural knowledge, wherein learners understand the rule theoretically but do not consistently apply it in practice.

For certain patterns, the error rate within the pattern was moderate, but its overall occurrence across all patterns remained low (G8(C) *(Verb + D O +for + I O =Verb + I O + D O), G8(F) (Verb + bare infinitive),* G8(D)/(d) *(Verb + (Object) + Preposition + Object)* and G3 *(noun + that clause)*). This suggests that the issue was specific to the individual pattern rather than a widespread problem affecting learners' overall grammatical accuracy. Among these patterns, the highest frequency was observed within pattern G8(C) *(Verb + D O +for + I O =Verb + I O + D O).* Within this pattern, 1 occurrence of *incomplete rule application* out of 4 (25%) might underscore intermittent deficiencies in the comprehensive implementation of this structure, while its overall frequency across all errors was 0.88%. This may imply partial mastery when learners comprehend the structure but inconsistently apply it. In the context, "*They prepare a lesson plan their students…",* the learner neglected the essential preposition '*for*', leading

to a grammatically incomplete phrase. This may indicate a limited comprehension of the G8(C) pattern, wherein learners acknowledge the necessity of a direct and indirect object but neglect to incorporate the preposition that connects them. Brown (2000) ascribed these errors to insufficient procedural knowledge, wherein learners understand the rule but fail to apply it consistently and comprehensively. Additionally, amidst these patterns, G3 *(noun + that clause)* had the lowest proportion. Within pattern G3, three examples of *incomplete rule application* out of 15 (20%) indicate that incomplete application of rules was a quite notable issue in the G3 pattern, while its overall frequency across all errors was 2,63%. These errors likely pointed to a persistent problem in which students were unable to fully form *noun + that clause* structures. This is consistent with research in the literature, including an empirical investigation on morphological errors made by Mongolian learners conducted by Bao (2015). Bao indicates that learners' difficulties with intralingual elements such as overgeneralisation and rule application are reflected in incomplete application errors, which he blames on a lack of core knowledge of English grammar. In the following sentence from the learner corpora, "*The idea that students should study regularly and complete…*", the student neglected to provide the requisite predicate, leading to an incomplete sentence. This can be ascribed to the fact that the learner may comprehend the necessity of a "*that-clause*" yet overlook the subsequent linking parts required to complete the sentence (Bao, 2015).

The lowest percentages occurred in patterns G5 *(adjective + preposition)* and G1 *(noun + preposition),* indicating that *incomplete application of rules* was a minor but present issue in these patterns, while their overall frequencies were quite low. In pattern G5 *(adjective + preposition),* errors were seen in 2 cases out of 30 (6,67%), with an overall rate of 1,75%. These errors likely arose from the incomplete application of a grammatical or syntactic rule, frequently resulting in omissions of prepositions or inconsistencies. In the instance, "*Being aware the media's influence is essential for developing…*", the absence of '*of*' signifies an unawareness that the adjective '*aware*' necessitates the preposition *'of'*. This inaccuracy may underscore learners' insufficient comprehension of *adjective-preposition* collocations, which are frequently unpredictable and must be acquired individually. Regarding pattern G1 *(noun + preposition)*, *incomplete applications of rules* occurred in 4 instances out of 65 (6,15%), with an overall rate of 3,51%. These errors mostly involved neglecting prepositions mandated by specific

English constructions. In the example, "*This can result from deficiency the personality*...", the student neglected to put the preposition '*in*' after the noun '*deficiency*'. Students might have overlooked or misunderstood the importance of associating prepositions such as '*in*' with specific nouns, indicating a limited proficiency in *noun + preposition* pattern.

The absence of incomplete application of rules errors across patterns G6, G8(J), G8(K), G8(L), G8(M), G8(N), G8(O), G8(P), and G8(R) indicated that in contrast to irregular structures, certain collocational patterns adhere to consistent formations, which minimises confusion. Additionally, the errors might not stem from their incomplete understanding of these structures, as these errors frequently arise when learners possess a partial comprehension of a rule but lack the necessary practice to consistently apply it (Brown, 2000).

5.3.3.5 Ignorance of rule restriction errors

Across all patterns, the total number of errors attributed to *ignorance of rule restrictions* revealed that this error source was most frequently observed in G8(D)/(d) (*Verb + (Object) + Preposition + Object*), accounting for 28.67% of the cases (41 out of 143). The second highest proportion was found in G4 (*Preposition + Noun*), comprising 20.28% of the total (29 out of 143). In contrast, the lowest occurrence rates were identified in patterns G8(A) (*Verb + Direct Object + to + Indirect Object = Verb + Indirect Object + Direct Object*) and G8(B) (*Verb + Direct Object + to + Indirect Object*), each with only one instance (0.70%) out of 143 (see Table 4.8, p. 71). This shows that *ignorance of rule restriction* errors was distributed more frequently among other patterns and may suggest that learners do not struggle as much with these patterns, indicating that some structures are more intuitively understood than others.

Regarding the prevalence of *ignorance of rule restriction* errors specifically within the patterns, the results revealed that the highest percentages of errors were seen in G4 *(preposition + noun)* and G8(R) *(It+ Verb + Object + to Infinitive/that-clause)* structures (See Table 4.10, p. 74). In G4 *(preposition + noun),* there was a dominance of this source of error with 29 out of 45 cases (64,44%), and the overall rate was 20,28%. This indicates that *ignorance of rule restriction* errors were dominant in the G4 pattern, highlighting that learners may have found it challenging to identify and utilise fixed prepositional

phrases, such as '*on the contrary*' in the following sentence, "*Many thought the new policy would decrease productivity; in the contrary, it led to significant improvements.*", the student seemed to erroneously employ '*in the contrary*' instead of the accurate phrase '*on the contrary*'. This shows that learners might lack awareness of the particular prepositional rules governing certain phrases. As stated by Liao and Fukuya (2004), this phenomenon is particularly prevalent with prepositions, as their application is frequently colloquial and differs throughout languages. This finding is consistent with Nacey and Graedler's (2015) research on how Norwegian EFL learners used prepositions in both written and spoken contexts. According to their study, learners frequently have trouble with the congruence of L1 and L2 prepositional structures, which results in the use of prepositions that are inappropriate for the context. Nevertheless, these results are in opposition to those of Neumanová (2023), who highlighted that L1 interference affects errors dominantly in various proficiency levels. This discrepancy can be attributed to the structural differences between the participants' native language (L1) and English (L2), which significantly influence prepositional errors. For instance, Norwegian and English may share more similarities in prepositional structures compared to Czech and English. As a result, the types and frequency of errors can vary depending on how prepositions function in the learners' first language.

Ignorance of rule restriction error occurred in 5 cases out 12 (41,67%) within G8(R) *(It+ Verb + Object + to Infinitive/that-clause)*, making it the most dominant source of error with an overall rate of 3,5%. This reveals that *ignorance of rule restriction* errors seemed to be a major issue, while this type of error was more prominent in other patterns. Learners frequently misapplied grammatical rules when constructing this structure, which might signify an unawareness of the rule. In the provided case, "*Lack of communication often leads to misunderstandings, and it causes to break up…*", the student neglected to supply the necessary noun or pronoun preceding '*to break up*', prohibiting the use of '*causes*' without a specified subject (e.g., causes + noun/pronoun + to infinitive). James (1998) observed that learners' failure to recognise rule restrictions often results from an incomplete understanding of the contextual and structural conditions that determine the applicability of particular grammatical rules, leading to misapplication in production.

Certain patterns exhibited a high prevalence of this error, making it the second most frequently occurring error type (G1 *(noun + preposition),* G8(E) *(Verb + to + Infinitive)*

and G5 *(adjective + preposition))*. In pattern G1 *(noun + preposition),* there was also a notable impact of *ignorance of rule restriction* with 20 instances (30,77%), with an overall rate of 13,99%, indicating that *ignorance of rule restriction* errors seemed to be an issue in this pattern due to a lack of understanding of grammatical restrictions. This also showed that learners were likely unaware of the fixed or restricted character of specific prepositional combinations with nouns. In the following example, "*They are deterrent factors with a crime*," the learner erroneously employed '*with*' instead of '*for*' following '*factors*', perhaps due to an inadequate understanding of the precise collocational demands of '*deterrent factors*'. This can be attributed to the fact that prepositional collocations frequently lack regular norms, making them difficult for learners to recall and use correctly (Swan, 2005).

In pattern G8(E) *(Verb + to + Infinitive)*, 13 instances out of 45 (28,89%) were *ignorance of rule restriction* errors, making it the second most prominent in this pattern. This shows that this type of error was a notable issue in the G8(E) pattern, while its overall frequency across all errors was lower, 9,09%. These errors mainly occurred when learners utilised improper prepositions or forms in contexts that demanded rigorous collocational adherence. To illustrate, the student mistakenly used '*for being*' in place of '*to be*' in the following sentence: "*The criminal pretended to be someone else to escape…*" This error may reflect a lack of understanding of the grammatical conventions that govern specific verb–preposition combinations in English. This observation is consistent with Nesselhauf (2005), who highlighted the frequent occurrence of such errors in collocational contexts, particularly among non-native speakers, attributing them to limited exposure to fixed lexical patterns and insufficient explicit instruction.

In pattern G5 *(adjective + preposition), ignorance of rule restriction* errors were observed as the second most prevalent error category with 8 instances out of 30 (26,67%). This suggests that this error was a major problem in the G5 pattern, and its overall frequency across all errors was 5.59%, which was relatively moderate. Learners frequently misused, omitted, or incorrectly applied prepositions after adjectives, possibly due to a lack of understanding of grammatical restrictions. In the instance, "*They may feel jealous with their posts and..",* the learner employed the erroneous preposition '*with*' instead of '*of*'. This inaccuracy indicates that the student seemed to be oblivious to the stringent collocational requirement that '*jealous*' is paired with '*of*' in English. As asserted by

Schmitt (2000), collocations provide a considerable barrier for learners due to their lack of adherence to predictable standards. Ellis (2003) further highlighted that learners not exposed to enough idiomatic collocations in the target language are likely unaware of the limitations.

In pattern G8(D)/(d) *(Verb + (Object) + Preposition + Object), ignorance of rule restriction* errors were observed in 41 cases out of 184 (22,28%) with an overall rate of 28,67%. The high frequency of these errors in the G8(D)/(d) pattern suggests that learners frequently misapplied grammatical rules; their errors may have arisen when learners did not identify the particular prepositions mandated by specific verbs. In the example, "*The residents began to complain of this situation…*", the verb *complain* should be followed by the preposition *about*, making the correct collocation *complain about something*. The learner's substitution of '*of*' in place of *about* reflects a failure to observe collocational restrictions, conventional prepositional pairings that are largely arbitrary and must be learned through exposure. This type of error suggests a limited familiarity with the lexical and grammatical constraints that govern verb-preposition combinations in English. As Richards and Schmidt (2013) note, such errors often arise when learners have not received sufficient input to internalise idiomatic and formulaic language use. The misuse of prepositions in collocational contexts is particularly common among EFL learners, as it requires not only knowledge of meaning but also sensitivity to native-like patterns that are typically acquired through frequent and meaningful exposure to authentic language use.

Some patterns had a moderate occurrence of this error both within the pattern itself and lower in overall terms. This indicated that the error was a moderate issue in these patterns; however, it had a minimal impact overall (G8*(Q) (Verb + (Object) + wh- clause/ wh-phrase), G8(L) (Verb + that-clause),* G8(N) *(Verb + Object + Complement), G8(B) (Verb + D O + to + I O),* G8(I) *(Verb + Object + Infinitive)).* Among these, the highest frequency was observed in G8(Q) *(Verb + (Object) + wh- clause/ wh-phrase)* with 3 cases out 20 (15%) within the pattern, while its overall frequency across all errors was 2,10%. This reveals that learners occasionally misapplied grammatical rules when using wh-clauses or wh-phrases as complements to verbs, possibly signifying an absence of comprehension regarding the syntactic constraints that regulate *wh-clauses,* leading to incorrect sentence construction. These errors often involved incorrect word order,

omission of necessary elements, or misplacement of *wh-clauses* within the sentence. In the given sample, "*Doctors asks them that they are against taking the medication…*", the learner might not have acknowledged that specific verbs necessitate *wh-words* (e.g., whether) instead of broad conjunctions such as '*that*'. Students may excessively utilise '*that*' as a universal connector because of its prevalent use across different clause types, indicating a restricted comprehension of collocations and syntactic rules, as Brown (2000) explains, often resulting from inadequate exposure to varied language structures.

Regarding pattern G8(L) *(Verb + that-clause), ignorance of rule restriction* errors arose in 2 cases out of 14, 14,29%. This indicates that this source of error was a moderate issue in the G8(L) pattern, while its overall frequency across all errors was 1.40%. These errors involved omitting necessary elements, misplacing the '*that'* conjunction, or applying incorrect verb complementation. In the sentence, "*They assert which there is no need for additional funding…"*, the learner used an incorrect conjunction, '*which*', instead of '*that*', likely stemming from uncertainty over their syntactic functions. This may show a lack of comprehension of rule restrictions, in which students are unaware of the rules that apply to the use of particular connections. In this case, it might have led them to treat *which* and *that* as interchangeable.

Regarding pattern G8(N) *(Verb + Object + Complement), ignorance of rule restriction* errors were observed in 3 cases out of 22 (13,64%), indicating that this type of error was a moderate issue in the G8(N) pattern, whereas its overall frequency across all errors was lower, 2.10%. These errors often involved misusing prepositions, incorrect word order, or applying an incorrect verb complementation structure. In the sentence, "*Making courses for compulsory ensures that all students gain essential knowledge and skills,*" the learner erroneously employed the preposition '*for*' in contexts where it was superfluous, perhaps stemming from a misunderstanding of the constraints on prepositional usage in English. Learners may have excessively generalised the necessity for prepositions in *Verb + Object + Complement* constructions, especially when articulating complex concepts. This is consistent with Aye's (2021) observation that regarding the G8(N) pattern, which necessitates that learners comprehend verb complementation and the syntactic function of complements in English, learners tend to ignore rules restricted to the use of this pattern.

Within pattern G8(B) *(Verb + D O + to + I O),* there was one instance of *ignorance of rule restrictions* error out of 8 (12,5%). This suggests that this source of error was a minor issue in the G8(B) pattern, occurring in only 1 out of 8 cases (12,5%), and its overall frequency across all errors was low (0,70%). The error in this pattern may indicate an issue with fixed expressions involving verbs and prepositions in the G8(B) pattern. The inaccuracy in the sentence, "*The use of digital media brings different dimensions at communication…*" might show that learners lack awareness of the fixed prepositional collocation '*bring dimensions to*'. The utilisation of '*at*' in place of '*to*' implies an unfamiliarity with this particular collocation. Richards and Schmidt (2013) assert that such errors highlight the challenges learners face in acquiring fixed expressions and prepositional patterns, as these are typically language-specific and not readily inferable from universal grammatical norms.

Some patterns had a lower occurrence of this error both within the pattern itself and in overall terms, highlighting that the error was not a significant issue in these patterns and had a minimal impact overall. To illustrate, within pattern G8(H) *(Verb + Object + to Infinitive), ignorance of rule restriction* errors occurred in 2 cases out of 22 (9,09%), while its overall frequency across all errors was 1,40%, indicating minimum misuse of the rigid characteristics of specific verb forms. In the sentence, "*Social media helps the people with connect with friends and family…*" the learner incorrectly used the preposition '*with*' in the situation that necessitated *to + infinitive*. Learners may erroneously associate '*with*' with causative relationships, resulting in mistakes such as '*with connect*'. This aligns with Swan (2005), who highlighted that learners frequently encounter difficulties with preposition usage in fixed expressions because of the discretionary nature of prepositional selections in English. Additionally, concerning the pattern G8(A) *(Verb + D O + to + I O = V + IO + D O)*, only one instance of *ignorance of rule restriction* error was identified within the pattern, accounting for 5.88% of errors (1 out of 18). Its overall occurrence across all patterns was minimal, comprising just 0.70% of the total errors. This indicates that *ignorance of rule restriction* errors was a minor issue in the G8(A) pattern. Only one recorded case of this source of error indicated a misunderstanding of the special rules regulating prepositions in this pattern. In the sentence, "*Local businesses offer a lot of savings of students on entertainment…*", the utilisation of '*of*' in place of '*to*' indicated that learners could not recognise the established

expression '*offer something to someone*' in English. This error can be attributed to the fact that collocation errors frequently arise from inadequate exposure to actual language, resulting in learners' inability to internalise the established connections among verbs, objects, and prepositions. Learners fail to grasp how words naturally fit together in a language (Nesselhauf, 2005).

Ignorance of rule restriction errors were not observed in patterns G2, G3, G6, G7, G8(C), G8(F), G8(G), G8(J), G8(K), G8(M), G8(O), and G8(S/s), possibly indicating internalisation of the correct applications of these structures. Typically, learners commit such errors by applying grammatical rules in contexts where they are not appropriate, which is frequently the result of inadequate learning or a failure to recognise existing structural constraints (Heydari and Beghari, 2012). The absence of these errors may indicate a comprehension of the rules of the target language, which reduces the likelihood of inaccuracies that arise from the improper application of these rules. This internalisation might be indicative of the successful acquisition of the language's structural nuances by learners, resulting in more accurate and fluent language use. This profound integration enables them to apply linguistic rules without conscious effort as their comprehension of grammar, syntax, and collocational patterns becomes more intuitive.

5.3.3.6 Approximation errors

Overall, the total amount of *approximation* errors across all patterns showed that this form of inaccuracy was primarily seen in *G8(D)/(d) (Verb + (Object) + Preposition + Object)* with a percentage of 53,33% (8 cases out of 15). The other six patterns had lower occurrences (G8(S/s)) (*Verb + Complement (Adj. or N)*) (2 out of 15), G1 (*noun + preposition*), G2 (*Noun + To + Infinitive*), G5 (*adjective + preposition*), G7 (*adjective + that clause*), and G8 (P) (*Verb + (Object) + Adverbial*) (1 out of 15 each)) (See Table 4.9, p. 72). According to this distribution, *approximation* was concentrated in certain grammatical areas rather than being a general problem.

When *approximation* errors were examined regarding the frequency of this error within each pattern, the highest rate was observed in pattern G7 *(adjective + that clause)* with 1 out of 7 instances (14,29%), making it the second most prominent error within this pattern (See Table 4.10, p. 74). The proportion of *approximation* for this pattern was 6,67% among all patterns. This indicates that these errors were rare overall, but present within

the G7 pattern. For instance, the error in the sentence, "*This situation seems undisputable…"* may indicate an effort to approach the proper word by replacing it with a phonetically similar yet non-standard term ('*undisputable*' for '*indisputable*'). This can be attributed to the fact that in order to address deficiencies in their lexical knowledge, learners frequently implement communication strategies that involve the use of related or newly developed forms to communicate meaning. For instance, they could describe an unfamiliar term or invent new terms to convey their ideas. These strategies allow learners to sustain communication flow in spite of vocabulary constraints (Putri, 2013).

Within the G2 pattern (*Noun + To + Infinitive*), there was also only 1 occurrence of *approximation* error out of 12 (8,33%), while the rate of it in this pattern accounted for 6.67% among all patterns. This indicates that *approximation* errors were rare within this pattern and that learners might have possessed adequate vocabulary for this structure but occasionally resorted to near substitutes in the absence of precise terminology. To illustrate, the learner used '*wanting*' instead of '*willingness*' in the following sentence: "*…due to their wanting to know more about this subject…*". In this case, the learners conflated '*willingness*' with '*wanting*', maybe owing to a restricted vocabulary or an insufficient comprehension of the subtle distinctions between the two concepts. The *approximation* error observed indicates that the learner substituted an incorrect structure that did not conform to standard *noun-to-infinitive* grammar. This is in line with the findings of Hong et al (2011), who investigated lexical collocation errors and found that *noun* errors are significantly affected by the strategy of *approximation*. This strategy reflects an attempt to compensate for limited vocabulary by relying on familiar terms, which can result in inaccuracies, emphasising the need for targeted vocabulary instruction to address these challenges.

In pattern G8(S/s) (*Verb + Complement (Adjective or Noun)*), *approximation* errors were identified in 2 out of 27 instances, representing 7.41% of the errors within this pattern. These accounted for 13.33% of all *approximation* errors across patterns. This suggests that such errors were relatively infrequent within the G8(S/s) structure, indicating that learners encountered fewer difficulties with *approximation* in this particular collocational pattern. Learners seemed to be in an attempt to construct comprehensive and grammatically correct sentences rather than simplifying or guessing structures. In the example, "*Children seem improductive if tasks are not assigned according to their*

interests," the learner used '*improductive'* instead of the correct term '*unproductive'*. This lexical error likely stems from reliance on phonological similarity, a common learner strategy when the accurate form is unknown or uncertain. Such errors may reflect attempts to apply productive morphological rules (e.g., prefixation) without fully grasping conventional usage, creating non-standard forms (Ringbom, 2007).

Within pattern G8(D)/(d) (*Verb + (Object) + Preposition + Object*), eight instances of approximation errors were identified, accounting for 4.35% of the errors in this pattern. Although the proportion of *approximation* errors in G8(D)/(d) was relatively low, the overall rate of approximation errors across all patterns was notably high at 53.33%. This suggests that while *approximation* was not a predominant issue within this specific pattern, it was more frequently associated with G8(D)/(d) than other grammatical structures when such errors occurred. Learners often substituted a preposition that was similar in meaning but did not fit the specific *verb-object-preposition* relationship. In the context, "*We should stop ourselves from praising actions that go against our values.*", '*stop*' was utilised in place of '*prevent*', leading to a little alteration in the intended meaning. This may be the result of a limited vocabulary or difficulty distinguishing between synonyms. Ellis (2003) noted that learners frequently approximate when they are uncertain about their vocabulary or grammar proficiency.

Within pattern G5 *(adjective + preposition),* there was only one case of *approximation* out of 30 (3,33%). The rate of approximation for this pattern was 6,67% among all patterns, indicating that the G5 pattern did not contribute to *approximation* errors highly, and its occurrence within the pattern was low. This suggests that learners were more likely to make *approximation* errors in other patterns rather than in *adjective-preposition* constructions. In the given example, "*The new policy became uneffective in reducing crime rates.*", the learner erroneously constructed '*uneffective*' instead of '*ineffective*', perhaps on account of confusion between the prefixes '*un-*' and '*in-*'. This may signify a limited understanding of morphology and an effort to estimate the target term using existing knowledge. *Approximation* errors are caused by learners' attempts to cover lexical gaps with similar-sounding or logically built forms, according to Richards and Schmidt (2013). Based on this, the infrequent occurrence of these errors implies that learners generally understand the basic structure of adjective-preposition pairings.

Nevertheless, they may occasionally overlook critical components or misuse prepositions due to their lack of experience.

The results showed that both the G8(P) *(Verb + (Object) + Adverbial)* and G1 *(Noun + Preposition)* patterns had lower rates of *approximation* errors. Only one approximation incident out of 56 cases (1.79%) was observed in G8(P), which accounted for 6.67% of total approximation mistakes across patterns. Likewise, in G1, the overall approximation error rate was 6.67%, with only one occurring out of 65 occurrences (1.54%). This indicates that learners committed relatively few total *approximation* errors, and errors in these patterns were more likely caused by other factors, such as L1 transfer or incomplete application of rules. When *approximation* errors did occur, they represented situations in which students used a phrase that was inaccurate or nonstandard yet roughly matched the intended meaning. To illustrate, while using the pattern G8(P) *(Verb + (Object) + Adverbial),* a learner demonstrated a misapprehension regarding English adverb formation. In the following instance, "*They are expected to act more thinkfully than others.*" '*thinkfully*' was used instead of '*thoughtfully*', indicative of incomplete understanding. As for the pattern G1 *(Noun + Preposition),* in the example provided, "*...feelings of inferiority and insatisfaction with her life…",* the learner replaced the correct term with an incorrect but similar word form, using '*insatisfaction*' instead of '*dissatisfaction*'. According to Hong et al. (2011), learners frequently employ the approximation strategy when the accurate structure in the target language has not yet been fully consolidated within their interlanguage system. This strategy functions as a compensatory mechanism, enabling learners to reduce their linguistic load while preserving communicative intent. By substituting an incorrect but semantically related form, learners attempt to maintain fluency despite gaps in grammatical or lexical knowledge.

Approximation error was not observed in patterns G3, G4, G6, G8(A), G8(B), G8(C), G8(E), G8(F), G8(G), G8(H), G8(I), G8(J), G8(K), G8(L), G8(M), G8(N), G8(O), G8(Q), and G8(R). The absence of this source of error in these patterns might indicate that learners attempted to conjecture or substitute words with near-equivalents when they used these structures. This suggests a more precise and accurate use of language, as learners utilised their consolidated knowledge of collocational structures rather than relying on *approximations* when confronted with gaps in their vocabulary.

Overall, the results revealed that learners' collocational challenges were impacted by mostly *intralingual* errors. This implies that errors were more likely to be the consequence of a lack of comprehension of English grammar than of direct interference from Turkish. Nevertheless, it was observed that learners' capacity to accurately employ English collocations was still influenced by negative transfer from Turkish. As a result, both intralingual factors and *L1 transfer* influenced learners' grammatical collocational use. Approximation errors were the least effective ones, and this might result from the nature of writing, which enables learners to plan, revise, and consult resources, thereby promoting greater accuracy.

Furthermore, as it was noted earlier, in the Turkish context the majority of previous research focused on prepositional patterns, leaving other grammatical collocation patterns unexplored, such as G8 (K) (*Verb + a possessive + V-ing*), G8 (M) (*Verb + Object + to be + Complement*) and G8 (R) (*It + Verb + Object + to Infinitive/that-claus*e). Therefore, the results of this investigation contribute to the effective teaching of collocations by exploring all grammatical patterns suggested by Benson et al. (2000). Additionally, the results highlighted the necessity of creating instructional strategies that are specifically designed to enhance the collocational proficiency of Turkish EFL students by mitigating L1 interference and addressing intralingual factors. The implementation of such targeted strategies can improve the overall proficiency of learners by promoting more accurate and fluent language use. Consequently, the results of this investigation may provide valuable insights for educators and researchers who are striving to enhance the quality of English language instruction in the Turkish context.

5.3 Suggestions

Drawing on the findings of the current study, this section presents implications (5.3.1) in light of earlier research on the subject. Limitations and suggestions for further studies are addressed in Section 5.3.2.

5.3.1 Implications

With the aim of addressing key issues raised by the findings, the implications will be discussed under two main headings: (1) methodological and technical implications, which address the design, analysis, and tools used in corpus-based research on collocations, and

(2) implications for textbooks and dictionaries, focusing on how the findings can inform the development of more effective language learning materials and reference resources.

5.3.1.1 Methodological and technical implications

The current research methodologically identified some considerations for language teaching. The primary implication of the current study is to allocate greater emphasis on collocation instruction. The study revealed the underutilisation and erroneous use of specific patterns such as G8 (Q) (*Verb + (Object) + WH-clause/WH-phrase*), G8 (R) pattern (*It + Verb + Object + to Infinitive/that-claus*e) G8 (K) (*Verb + a possessive + V-ing*) and G8 (M) (*Verb + Object + to be + Complement*) which were considered to be rather complex. Owing to the complexity and specialised nature of these patterns, they were less common in general language use, as supported by previous research (e.g. Boland et al., 1995; Bardovi-Harlig, 2002; Hunston, 2003; Collins, 2004; Yoon, 2008; Hassan, 2014). Learners tended to avoid such constructions that are less intuitive and more idiomatic; as a result, they frequently chose less complex, more straightforward sentence structures (Granger, 1998). As well as underuse the significant overuse of specific patterns, such as G8(F) (*verb + bare infinitive*), further reinforced the fact that learners tended to employ simpler, high-frequency forms when constructing sentences (Cobb, 2003; Laufer, 2003), consequently impeding academic sophistication and native-like fluency (Larsson, 2012). Larsen-Freeman (2014) attributes the challenge that learners face when they deal with collocations to a lack of natural language in formal education. This underscores the fact that explicit teaching of complex collocations to EFL learners is essential to help them understand and internalise the fixed or semi-fixed relationships between words, their meanings, and their use in context in order to achieve academic mastery and fluency comparable to those of native speakers (Lewis, 2000; Ellis, 2008).

Furthermore, although learners tended to avoid complex structures—particularly G8(R) (*It + Verb + Object + to-Infinitive/that-clause*), G8(K) (*Verb + a possessive + V-ing*), and G8(M) (*Verb + Object + to be + Complement*), the error rates within these patterns remained relatively high, indicating that these structures posed considerable challenges for learners. Further considerations for teaching can be derived from the fact that some patterns indicated a challenge for learners, such as prepositional structures (e.g., G1 (*Noun + Preposition*), G8 (D/d) (*Verb + (Object) + Preposition + Object*)), which showed a moderate to high frequency of errors. The learners' errors primarily stemmed

from *ignorance of rule restriction* and *transfer* from their first language (L1). Consequently, these findings emphasise the necessity of focused instruction that addresses these patterns and error types during formal and academic instruction to improve learners' syntactic and discourse-level competence (Chen, 2024).

In light of the aforementioned issues, methodological approaches in ESL/EFL contexts need to place greater emphasis on grammatical collocations to add more variety and accuracy to their production. This underscores the necessity of explicit teaching strategies, as mentioned earlier, to acquaint learners with the expressive and functional potential of certain patterns, thereby facilitating the development of more nuanced and idiomatic language production, as suggested by Marco (1999). Teachers can assist EFL learners in mastering collocations by employing techniques like repetition, corpus tools, and L1 contrast, which will improve their fluency, accuracy, and self-assurance when speaking English (Nation, 2001; McCarthy and O'Dell, 2005).

For EFL/ESL learners, effective teaching of collocations necessitates a methodical approach that combines exposure to real language and meaningful practice (Boers et al., 2006; Lewis, 2000). Due to the non-intuitive nature of collocations, particularly complex ones, which cannot be easily deduced from the meanings of individual words, learners often encounter difficulties with them. Explicit instruction of collocations, teaching them as meaningful components rather than isolated words, enhances students' comprehension of their fixed nature and appropriate use, as indicated by Lewis (1997). Therefore, targeted lessons can clarify the meaning, grammatical structure, and usage of complicated collocations (such as *take responsibility for, on the edge of*). Secondly, for teaching collocations, Thornbury (2002) promotes contextualised information, pointing out that word retention and comprehension are enhanced when encountered in relevant circumstances. Students can encounter collocations in significant contexts through authentic materials such as articles, dialogues, or writings that inherently include complex collocations and highlight them for observation and reflection or discuss their function within the context (e.g., *exert maximum effort in a collaborative article*). Third, for students to internalise the structure and usage of complex collocations, they must be exposed to them repeatedly. Nation (2001) emphasises the value of repetition in vocabulary learning, stressing that in order for students to properly internalise a word or phrase, they must encounter it several times. Employing spaced repetition strategies, like

gradually adding target collocations to speaking, writing, and comprehension exercises, would be beneficial. Fourth, in order to balance the development of accuracy and fluency in vocabulary learning, Ellis (2009) advocates combining both controlled and unrestricted practice. Accurate and fluid use of complex collocations can be made possible by activities under control, including matching collocations, fill-in-the-blank exercises, and sentence completion assignments. Another method, as suggested by Ellis (2009), is corrective feedback which would be beneficial for encouraging proper language use by giving students targeted comments on collocation mistakes to help them improve their comprehension and application. One way is to address learners' mistakes in complicated collocations by using corrective feedback strategies such as recasts and metalinguistic explanations. Additionally, Wray (2002) suggests paying attention to formulaic words as mastering formulaic sequences improves fluency and lessens the mental strain associated with language output. One method might be to use colloquial collocations and popular formulaic statements (e.g. *on the brink of disaster*) as pre-made language units.

As for overcoming *L1 negative transfer* errors, comparing and contrasting similarities and differences between L1 and L2 collocations might help EFL learners attain higher collocation proficiency by increasing learners' awareness of the concept of L2 collocation and its characteristics by exposing learners to various types of collocations in L2 classroom settings (Özata, 2020). As Odlin (1989) noted, it is crucial to be aware of the ways to handle cross-linguistic variations in language instruction. Concerning the *intralingual* errors prevalent in this study, they may require targeted instructional strategies focusing on learners' understanding of grammatical rules and their appropriate application. Overcoming issues like overgeneralisation and ignorance of rule restrictions necessitates exposure to diverse and authentic language input, as well as explicit teaching of collocational patterns (Sawai et al. (2013). In this way, as Shin and Nation (2007) emphasise, L2 learners' English may sound more natural.

Another technique for teaching collocations is incorporating academic writing into ESL/EFL contexts, which is consistent with language acquisition and pedagogical efficacy studies. Engaging learners with authentic, contextually relevant uses of language via academic writing not only improves their linguistic skills but also equips them for effective academic and professional communication (Cortes, 2004). Empirical research underscores the advantages of integrating collocation education into academic writing.

Boers and Lindstromberg (2008) discovered that highlighting collocations in writing assignments markedly enhanced learners' recall and their capacity to utilise collocations in novel contexts. What is more, corpus studies indicate that academic writing employs a more specialised and limited vocabulary than conversational English, with numerous specialised terms manifesting as fixed collocations (Hyland, 2004). Additionally, incontrast to passive recognition, academic writing necessitates that learners actively generate language, hence emphasising productive linguistic abilities. As stated by Nation (2001), this active involvement strengthens the proper utilisation of collocations. Nation (2001) further emphasises the importance of using lexical resources to improve students' comprehension of word combinations; therefore, an effective teaching strategy for writing courses would be to use collocation dictionaries and tools. Using tools like the Oxford Collocations Dictionary during academic writing tasks can help students identify and practice appropriate collocations. In addition, explicitly teaching collocations as part of academic writing instruction can significantly improve learners' proficiency (Lewis, 2000). For example, writing tasks can be designed to include target collocations, and teachers can provide feedback on collocation errors. In order to help reinforce correct usage, another strategy would be to engage in peer review sessions where students critique each other's use of collocations in academic writing, complemented by teacher feedback on collocation errors, which is also crucial for refining learners' written output, as suggested by Lewis (2000).

Integrating corpus resources into collocation instruction can also be effective as it enhances learners' comprehension of word combinations and fosters vital abilities for autonomous, data-driven language acquisition. As highlighted by Hunston (2002), this method improves the understanding of natural language usage since corpora offer authentic instances of collocational patterns, contextual subtleties, and word frequencies, assisting learners in prioritising prevalent combinations and circumventing errors. This method fosters contextual learning, error correction, and independent inquiry, allowing learners to contrast their usage with authentic facts. Practical applications encompass guided analysis, discovery learning, and writing assistance, rendering corpus-based learning an effective instrument for mastering collocations and connecting classroom instruction with authentic communication (Hunston, 2002).

As a subset of corpus-based learning in order to increase the competency of collocations, recent studies have emphasised the potential of Data-Driven Learning (DDL), which involves learners working directly with corpus data to facilitate language acquisition in various educational contexts. Explicit DDL instruction has been demonstrated to improve the collocational ability of EFL/ESL learners markedly. According to the findings of related research, students exposed to DDL methods exhibited significant enhancements in confidence and retention of collocational information, attaining superior results in both receptive and productive skills relative to those employing conventional teaching methods (e.g. Özbay and Özer, 2017; Liu, 2020; Alruwaili, 2020; Bernal, 2022; Liu and Gablasova, 2023). As suggested by Sun and Park (2023), by integrating authentic language data, the corpus offers a realistic learning environment, hence enhancing the authenticity of the educational experience. Providing learners with collocation-rich datasets from corpus tools allows them to discern the usage of collocations by native speakers across various contexts (Cortes, 2004). This genuine input fosters fluency and aids natural language acquisition by improving learners' comprehension of collocations in diverse real-world contexts (Nation, 2001). Furthermore, corpus-based DDL enhances learners' awareness of the target language and increases their exposure to genuine texts via concordance. It is accessible independently by learners, elevates motivation, and fosters learner autonomy (Sun and Park, 2023).

As the pioneer of DDL, Johns (1991) noted that learner-centred discovery entails providing learners with authentic corpus data, enabling them to observe and infer collocation principles through practical instances. To illustrate, by utilising corpus tools such as COCA (Corpus of Contemporary American English) or BNC (British National Corpus), one can investigate specific collocations, examine their application in sentences, and analyse the contextual framework, thus aiding learners in discerning their appropriate application. In addition, error correction can be effectively accomplished using real data by contrasting learner-generated phrases with corpus examples to identify errors. For instance, learners can ascertain that '*in purpose*' is erroneous by analysing corpus data that consistently illustrates the proper form, '*on purpose*' (Hunston, 2002). As a consequence, learners can enhance their understanding of linguistic structures and refine their language acquisition process by actively identifying and analysing discrepancies between their native language (L1) and the target language (L2). This recognition allows

them to rectify errors that result from L1 transfer, which occurs when the native language influences the usage of L2 in an inaccurate manner, as well as errors that arise from overgeneralisation, which occurs when learners mistakenly apply broad language rules to contexts that necessitate more nuanced applications or exceptions. Learners can progressively internalise correct forms, improve their accuracy, and develop confidence in their language proficiency through this reflective process.

5.3.1.2 Implications for textbooks and dictionaries

Textbooks have often been criticised for their insufficient treatment of collocations, particularly in terms of precision and contextual nuance (e.g. Peksoy, 2013; Ergül, 2014; Skopinskaja and Liiv, 2015). However, effective grammatical collocation training necessitates the integration of practical exercises, contextual examples, and explicit instruction in textbooks. Research consistently shows that tailoring collocation teaching to address learners' specific needs considerably enhances their language proficiency. Eid and Al-Jamal (2023), for instance, stress how crucial it is to incorporate collocations into communicative contexts - like essays and dialogues - in order to give learning purpose. Furthermore, teaching collocations along with their syntactic and semantic contexts aids students in comprehending how these pairings naturally occur in language use. Therefore, a key strategy for effective collocation teaching is explicitly giving grammatical collocations and their function in language use in textbooks. For example, teaching learners the grammatical dependencies of collocations like '*It surprised me to learn …*' and '*the foresight to do*' aids in their understanding of how to use them accurately. In addition, understanding is improved when collocations are included in relevant settings. For instance, learners can observe how collocations work in natural communication when complete sentences are used rather than separate words. Using real-world examples from sources such as news stories or narratives helps close the gap between classroom instruction and everyday language use. Another method is using webs of collocations in textbooks, as asserted by Handl (2009). Collocational webs are visual instruments that outline the interconnections among collocations, thereby elucidating these connections for enhanced comprehension. For example, a lexical network for the term '*take*' might encompass expressions like '*take responsibility*', '*take action*' and '*take a chance*'. These graphic representations enhance understanding and retention by depicting the relationships among collocations, hence assisting learners in comprehending their

contextual application more efficiently. Additionally, to help students avoid errors, textbooks can address common learner errors in the use of collocations. As Handl (2009) suggests, incorporating exercises such as error correction serves to further emphasise the appropriate use of collocations. For instance, pointing out mistakes like using '*interested on*' instead of '*interested in*' and providing an explanation might help students use the words correctly. Furthermore, for active learning of collocations, textbooks should include engaging activities and hands-on exercises, including gap-filling, sentence completion, matching, and translation. By encouraging students to use collocations to build phrases, these exercises may help them become more accurate and fluent. According to Petkovska and Neshkovska (2019), these kinds of assignments are important for encouraging participation and making sure that students internalise the right collocations through regular practice. Lastly, textbooks can emphasise nuanced variations since collocations frequently have pragmatic and cultural importance. For instance, the formality or tone of '*pay attention to*' and '*take notice of*' can change according to the situation. Learners can make language choices that are appropriate for the context by attending to these subtleties. All in all, textbooks should incorporate explicit instruction, offer a variety of contextualised examples, and incorporate exercises for real-world application in order to teach grammatical collocations. Learners' comprehension and retention can be further improved by correcting common mistakes and incorporating resources like collocational webs. Textbooks can make learning more relevant and meaningful by incorporating collocations into real-world communicative contexts, which will ultimately increase students' fluency and competency.

As essential resources for language acquisition, dictionaries are also essential for teaching grammatical collocations, as they enable students to comprehend, identify, and employ these structures efficiently. Effective collocational entry arrangement is crucial in electronic learner dictionaries, according to Lan and Yang (2020). According to them, learners' capacity to consult and employ collocations correctly is enhanced by clear organisation and real-world examples, accompanied by syntactic explanations of verb forms or prepositions to make the underlying grammar more understandable. Furthermore, in addition to their application in textbooks, Handl (2009) proposes that collocational webs should be integrated into dictionaries to enable students to quickly recognise patterns and connections. Another technique is to give students examples in

context to aid in their understanding of how collocations are used in everyday speech. In order to ensure that learners understand the semantic and pragmatic nuances of collocations, Torner and Arias-Badia (2023) emphasise the necessity of lexicographic tools that convey them in a variety of settings. Adding interactive features to digital dictionaries could be another technique. To strengthen collocation learning, digital dictionaries might include interactive features like tests, fill-in-the-blank activities, and pronunciation guides. Features like guided dictionary tasks not only strengthen learners' autonomy but also enhance their competence in using collocations accurately and contextually (Kim, 2017). Lastly, specialised and multilingual dictionaries can be utilised. By analysing collocation usage across languages, multilingual dictionaries, such as those covered by the Dictionaries of LSP Collocations (2022) project, address the demands of bilingual or professional users. All in all, dictionaries can be improved as teaching aids for grammatical collocations by adding contextualised examples, interactive features, error analysis, and comprehensive information. In addition to improving learning outcomes, these enhancements can help to promote deeper linguistic competency, allowing students to employ collocations correctly and smoothly in everyday situations.

5.3.2 Limitations and suggestions for further studies

One notable limitation of the present study concerns the differing time periods in which the essays from native and non-native participants were written. The reference corpus, drawn from a subset of LOCNESS, consists of argumentative essays authored by American and British university students, most of which date back to 1995. In contrast, the non-native learner corpus comprises essays written between 2009 and 2023. This temporal gap may have influenced language use, as linguistic norms and usage patterns can evolve over time, potentially affecting the comparability of the two datasets. This time gap introduces several potential issues: language usage and writing standards undergo evolution, indicating that essays composed in 1995 may not effectively represent contemporary collocational preferences or writing styles. Moreover, educational standards, pedagogical approaches, and evaluative processes have presumably evolved throughout time, potentially influencing the manner in which learners generate texts. Furthermore, the enhanced accessibility of internet resources, language-learning tools, and writing aids in recent years offers non-native learners benefits that were previously less accessible or altogether inaccessible to native writers in prior times. Consequently,

any discrepancies observed between the native and non-native corpora may not just indicate proficiency-related concerns but may also be affected by shifts in language usage, progressive educational methodologies, and technology innovations over time. This temporal discrepancy may restrict the validity of direct comparisons between the two groups.

Moreover, it is important to acknowledge that the corpora used in this study were collected within instructional settings. The written texts, produced by both native and non-native participants, were composed in response to specific prompt words. In this context, writers, regardless of their linguistic background, may have focused more on demonstrating writing proficiency than on expressing their authentic opinions. This dynamic could potentially influence both the frequency and accuracy of the collocational patterns observed in the data.

The analysis in this study was singularly focused on errors discovered in a learner corpus. The first suggestion for future research could be to broaden the scope by comparing errors in both native and learner corpora, thereby further enriching the comprehension of grammatical collocation usage. This comparative analysis has the potential to provide numerous significant insights regarding the frequency of errors, the types of errors, the impact of proficiency, and the cross-linguistic influence.

Additionally, collocational usage at varying proficiency levels was beyond the scope of this study. Therefore, future studies may examine collocation usage across several proficiency levels—beginner, intermediate, and advanced learners—to attain a more refined comprehension of the evolution of collocational competence over time. This comparison method may uncover significant distinctions in the ways learners at different levels of language acquisition utilise, comprehend, and internalise collocations.

Another suggestion involves incorporating qualitative methods. In the error analysis section, as a characteristic of corpus-based investigations, it was not feasible to confirm the sources of errors by enquiring directly with each learner for each example. Therefore, a qualitative follow-up study can be designed with a select group of specific learners, allowing for the exploration of their cognitive strategies through individual interviews or think-aloud protocols regarding collocation usage.

A further future study could also concentrate on the analysis of the utilisation of particles within these structures by examining specific patterns of grammatical collocations. For example, the *noun + preposition* pattern could be used as a focused framework to investigate the use of prepositions in collocations, with a particular emphasis on the choices of both native speakers and learners. In this way, common patterns can be identified, or native and learner use can be compared.

Finally, experimental investigations examining various instructional methodologies may yield significant insights into the most effective means by which learners acquire grammatical collocations. Researchers can generate evidence-based recommendations for enhancing collocation instruction in classroom environments by analysing diverse teaching methodologies and assessing their respective strengths and limitations. To illustrate, a study could compare the efficacy of explicit instruction, task-based learning, and data-driven learning to ascertain which method yields the best results for learners.

REFERENCES

Ahmed, M. B. (2008). Collocation Errors Made by Iraqi EFL Learners at University Level. *Tikrit University Journal of Humanities, 15*(11), 55-73.

Akıncı, M. (2009). *Effectiveness of corpus consultation in teaching verb-noun collocations to first-year ELT students.* Boğaziçi University.

Akkoyunlu, A. N. (2017). *Teaching collocations through data-driven learning activities in a translation course for prospective ELT teachers*. Çukurova University.

Aksungur, M. T. (2021). *The explicit teaching of L2 English collocations through graphic organisers*. Yeditepe University.

Alan, D. (2015). False concept hypothesised: Understanding misinterpretations in second language acquisition. *Journal of Language Learning, 22*(3), 134–158.

Al-Halalmeh, B. A. (2011). *An investigation of grammatical collocations among English majors at Mu'tah University*. Mu'tah University, Jordan.

Alhaysony, M. H. (2017). Strategies and difficulties of understanding English idioms: A case study of Saudi University EFL students. *International Journal of English Linguistics, 7*(3), 70–84. https://doi.org/10.5539/ijel.v7n3p70

Allen, J. P. and Corder, S. P. (1974). *Techniques in applied linguistics*. London and New York: Oxford University Press.

Alqarni, A. **S.** (2019). *Applying corpus pattern analysis to learner corpora: Investigating the pedagogical potential of the Pattern Dictionary of English Verbs* [Doctoral dissertation, University of Birmingham]. eTheses Repository. https://etheses.bham.ac.uk/id/eprint/9300/

Al-Sobhi, B. M. S. (2009). L1 interference in Arab ESL secondary school students' spelling errors. *International Journal of Education and Literacy Studies, 7*(3), 45–60. https://www.aiac.org.au/journals/index.php/IJELS/article/view/3735

Alsulayyi, M. N. (2015). The Use of Grammatical Collocations by Advanced Saudi EFL

Learners in the UK and KSA. *International Journal of English Linguistics; 5*(1), 32-43.

Al-Tamimi, A. (2006). Gradation in training materials: A factor in learner misinterpretation of grammar rules. *Second Language Studies, 14*(2), 89–112.

Altenberg, B., and Granger, S. (2001). The grammatical and lexical patterning of *make* in native and non-native student writing. *Applied Linguistics, 22*(2), 173–194.

Amara, N. (2015). Errors correction in foreign language teaching. *The Online Journal of New Horizons in Education*, *5*(3), 58-68.

Anthony, L. (2022). *TagAnt* (Version 2.0.5) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Anthony, L. (2023). *AntConc* (Version 4.2.4) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Ataç, İ., and Önal, A. (2022). First language influence in acquiring adjective-preposition collocations: A corpus-based analysis. *Linguistic Inquiry, 18*(3), 112–130.

Avcı, G. B. (2006). *Yazma çalışmaları ile desteklenen eşgörüntüleme öğretiminin öğrencilerin kelime edinimine etkileri*. Marmara University.

Ayar, Z. (2020). Error analysis of Turkish learners' English paragraphs from lexical and grammatical aspects. *ELT Research Journal, 9*(2), 34–50. https://dergipark.org.tr/en/download/article-file/1372135

Aydoğan, S. (2016). *Task orientation and vocabulary knowledge types: How do they relate to collocational knowledge?* Boğaziçi University.

Aye, S. (2021). Grammatical Collocation Errors made by Postgraduate English Specialization Students Studying at Yadanabon University. *MERAL Portal*. Retrieved from: https://meral.edu.mm/records/7402/export/json

Babanoğlu, M. P. (2011). The acquisition of English dative alternation by Turkish adult learners of English. *Dil Dergisi*, (153), 53-72.

Bağcı, N. D. (2014). *Turkish university-level EFL learners' collocational knowledge at receptive and productive levels.* Middle East Technical University.

Bahardoust, M. and Moeini, M.R. (2012). Lexical and Grammatical Collocations in Writing Production of EFL Learners. *The Journal of Applied Linguistics (5)*1, 62-86.

Bahns, J. (1993). Lexical collocation: a contrastive view. *ELT Journal, 47* (1), 56-63.

Bahns, J., and Eldaw, M. (1993). Should we teach EFL students collocations? *System, 21*(1), 101-114.

Balcı, Ö. (2006). *Teaching vocabulary through collocations and cliches in EFL classes.* Selçuk University.

Bao, X. (2015). *Empirical study on morphological errors of Mongolian learners in English writing*. *English Language and Literature Studies, 5*(3), 25–38. https://www.ccsenet.org/journal/index.php/ells/article/download/45798/24724

Bardovi-Harlig, K. (2002). A New Starting Point? Investigating Formulaic Sequences in Second Language Development. In R. Kaplan (Ed.), *The Oxford Handbook of Applied Linguistics* (pp. 154–163). Oxford: Oxford University Press.

Bartan, Ö. Ş. (2019). Lexical Collocation Errors in Literary Translation. *Dil Dergisi, 170*(1), 73-88. doi: 10.33690/dilder.528981

Başal, A. (2019). Learning collocations: Effects of online tools on teaching English adjective-noun collocations. *British Journal of Educational Technology, 50*(1), 342-356. https://doi.org/10.1111/bjet.12562

Benson, M., Benson, E. and Ilson, R. (1986). *The BBI Combinatory Dictionary of English*. John Benjamins Publishing Company.

Benson, M., Benson, E., and Ilson, R. (1990). *The BBI Dictionary of English Word Combinations.* Amsterdam/Philadelphia: Benjamins.

Benson, M., Benson, E., and Ilson, R. (1997). *The BBI Dictionary of English Word Combinations.* (Morton Benson, E. Benson, and R. Ilson, Eds.). John Benjamin Publishing Company.

Benson, M., Benson, E., and Ilson, R. (2010). *The BBI Dictionary of English: Your guide to collocations and grammar*. John Benjamins Publishing Company.

Bıçkı, A. (2012). *Acquisition of English collocations by adult Turkish L2 learners*. Çukurova University.

Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing, 8*(4), 243–257. https://doi.org/10.1093/llc/8.4.243

Biber, D. (2010). Corpus-based and corpus-driven analyses of language variation and use. In B. Heine and H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis* (pp. 159–192). Oxford: Oxford University Press.

Biber, D., Conrad, S., and Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. New York: Cambridge University Press.

Boers, F., and Lindstromberg, S. (2008). *Optimizing a lexical approach to instructed second language acquisition*. Palgrave Macmillan.

Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., and Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of memory and language*, *34*(6), 774-806.6.

Boonraksa, T. and Naisena, S. (2022). A Study on English Collocation Errors of Thai EFL Students. English Language Teaching. *Canadian Center of Science and Education, 15* (1), 164-177.

Borhan, M. M. (2024). Wh-phrases and their Interference of Interlingual in EFL Context in Afghanistan: EFL Context in Afghanistan. *International Journal of Curriculum and Instruction*, *16*(2), 275-286.

Bourke, J. M. (2007). Verbal complementation: A pedagogical challenge. *Reflections on English Language Teaching*, *6*(1), 35-50.

Bozoğlu, U. (2020). *A study on teaching collocations (Verb+noun) in a secondary school EFL context*. Muğla Sıtkı Koçman University.

Brown, D. F. (1974). Advanced vocabulary teaching: The problem of collocation. *RELC Journal, 5*(2), 1–11. https://doi.org/10.1177/00336882740050020

Brown, H. D. (2000). *Principles of Language Learning and Teaching* (4th ed.). Pearson Education.

Bui, T. L. (2021). The Role of Collocations in the English Teaching and Learning. *International Jourrnal of TESOL and Education, 1*(2), 99-109. Retrieved from https://eoi.citefactor.org/10.11250/ijte.01.02.006

Calude, A. S. (2017). Clausal Complements in Native and Learner Spoken English: A Corpus-Based Study with LINDSEI and VICOLSE. *Australian Journal of Linguistics*, *37*(1), 91–94. https://doi.org/10.1080/07268602.2015.1135707

Cambridge University Press. (n.d.). *Cambridge Dictionary*. Retrieved March 20, 2024, from https://dictionary.cambridge.org

Campoy, M. C., Belles, B., and Gea, M. L. (2010). Introduction to corpus linguistics and ELT. In M. C. Campoy, B. Belles, and M. L. Gea (Eds.), *Corpus-based approaches to English language teaching*. London: Continuum International Publishing Group.

Can, C. (2017). A Learner Corpus-Based Study on Verb Errors of Turkish EFL Learners. *Journal of Education and Training Studies*, *5*(9), 167-175.

Celce-Murcia, M., and Larsen-Freeman, D. (1999). *The Grammar Book: An ESL/EFL Teacher's Course (2nd ed.)*. Boston: Heinle and Heinle.

Chan, T. P., and Liou, H. C. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning*, 18(3), 231–250. https://doi.org/10.1080/09588220500185769

Chelli, S. (2013). *Interlingual and intralingual errors in the use of prepositions and articles.* Retrieved 2023, http://archives.univbiskra.dz/bitstream/123456789/3571/1/Interlingual

Chen, M. H. (2008). *A study of the English collocation competence of college students in Taiwan*. Unpublished master's thesis, National Taiwan University of Science and Technology.

Chen, X. (2024). Corpus-based study on language errors in English writing. *International Journal of Education and Humanities*, *15*(2), 144–153. https://doi.org/10.54097/f2cmct16

Cobb, T. (2003). Analysing late interlanguage with learner corpora: Quebec replications of three European studies. *Canadian Modern Language Review*, *59*(3), 393-424.

Collins, P. (2004). The Semantics of Complementation in English. *Journal of English Linguistics, 32*(1), 31–62. doi:10.1177/007542400403200103

Conrad, S. M. (1999). The importance of corpus-based research for language teachers. *System, 27*, 1–18.

Conrad, S., and Biber, D. (2001). *Variation in English: Multidimensional Studies.* London: Routledge.

Corder, S. P. (1981). *Error Analysis and Interlanguage*. Oxford University Press.

Cortes, V. (2004). Lexical bundles in academic writing: A corpus-based study of native and non-native English speakers. *Journal of English for Academic Purposes, 3*(4), 397–423.

Cowie, A. P. (1998). Past achievements and current trends. In A. P. Cowie (Ed.), *Phraseology: Theory, Analysis and Applications*. (pp. 1–20). Oxford: Clarendon Press.

Cowie, A. P., and Ronald, M. (1975). *Oxford dictionary of current idiomatic English*. Oxford University Press.

Cowie, A. P., Mackin, R., and McCaig, I. R. (1984). *Oxford Dictionary of Current Idiomatic English*. Oxford University Press.

Çelik, F., and Eren, Z. (2023). Taking advantage of L2 errors in EFL learners' written language production. *Journal of Language and Linguistic Studies, 19*(4), 78–92. https://www.researchgate.net/publication/375343011

Çepni, S. B. (2020). *Teaching collocations through data-driven learning: Comparison of two approaches.* Hacettepe University.

Dagneaux, E., Dennes, S., and Granger, S. (1998). Computer-aided error analysis. *System, 26,* 163-174.

Davies, M. (2008). *The Corpus of Contemporary American English (COCA): 600 million words, 1990-present.* Retrieved January 20, 2020, from https://www.english-corpora.org/coca/

DeCapua, A. (2008). *Grammar for Teachers: A Guide to American English for Native and Non-Native Speakers*. Springer.

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning, 55*(1), 1–25. https://doi.org/10.1111/j.0023-8333.2005.00294.x

Demir, C. (2016). *Lexical proficiency of collocation, boosting, and hedging in academic discourse: A comparative study*. Atatürk University.

Demir, C. (2018). Word combinations in academic writing. *Journal of Language and Linguistic Studies. 14*(1), 293-327.

Demir, S. (2022). *A comparative corpus-based analysis of lexical collocations in the EFL coursebooks*. Maltepe Üniversitesi.

Demirel, E. T. (2017). Detection of Common Errors in Turkish EFL Students' Writing through a Corpus Analytic Approach. *English Language Teaching*, *10*(10), 159-178.

Demirezen, M. (2021). The stress and intonation background of prepositional verbs versus phrasal verbs for language teachers. *Çankaya University Journal of Humanities and Social Sciences*, 15(2), 242–258.

Deshors, S. C. (2015). A multifactorial approach to gerundial and to-infinitival verb-complementation patterns in native and non-native English. *English Text Construction*, *8*(2), 207–235. https://doi.org/10.1075/ETC.8.2.04DES

Deshors, S. C., and Gries, S. T. (2016). Profiling verb complementation constructions across New Englishes: A two-step random forests analysis of ing vs. to complements. *International Journal of Corpus Linguistics*, *21*(2), 192-218.

Dinçer, A., Harmandar, B. B., and Sucak, D. (2022). An Analysis on Grammatical Collocations in Written Texts of EFL Learners. In *Proceedings of the First International Language-for-All Conference (LFAC'22)* (p. 31). Çukurova University. https://lfac.cu.edu.tr

Do, N. H. M., and Le, Q. T. (2023). Lexical Collocation Errors in Essay Writing: A Study into Vietnamese EFL Students and Their Perceptions. *International Journal of Language Instruction, 2*(2), 1-20.

Domede, A. B. (2024). Breaking the rules: Investigating grammatical errors in university writing. *Psychology and Education: A Multidisciplinary Journal, 61*(2), 123–140. https://scimatic.org/show_manuscript/2716

Doughty, C., and Williams, J. (1998). *Focus on form in classroom second language acquisition*. Cambridge University Press.

Dravishi, S., Hashemi, M. , and Azizinezhad, M. (2011). Collocation a neglected aspect in teaching and learning EFL. *Procedia - Social and Behavioral Sciences, 31*, 522–525.

Du X., Afzaal, M. and Al Fadda, H. (2022). Collocation Use in EFL Learners' Writing Across Multiple Language Proficiencies: A Corpus-Driven Study. *Front. Psychol*. 13:752134. doi:10.3389/fpsyg.2022.752134

Durmuş, B. Y. (2019). *The differential effect of explicit and implicit instruction on secondary school learners' collocational knowledge*. Çukurova University.

Durrant, P., and Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics in Language Teaching*, vol. 47, no. 2, 2009, pp. 157-177. https://doi.org/10.1515/iral.2009.007

Eid, M., and Al-Jamal, D. (2023). The role of collocations in communicative language learning: Integrating lexical patterns into essays and dialogues. *Journal of Language Teaching and Learning, 15*(2), 67–89.

Eker, S. (2001). *Development of collocational competence in a second language.* Uludağ University.

Ellis, N. C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, *24*(2), 143–188. https://doi.org/10.1017/S0272263102002024

Ellis, R. (1985). *Understanding Second Language Acquisition*. Oxford: Oxford University Press.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Ellis, R. (2006). Current Issues in the Teaching of Grammar: An SLA Perspective. *TESOL Quarterly, 40*(1), 83–107.

Ellis, R. (2008). *The study of second language acquisition.* Oxford: Oxford University Press.

Ergül, Y. (2014). *The effectiveness of using corpus-based materials in vocabulary teaching*. Unpublished MA Thesis. Denizli: Pamukkale University, The Institute of Educational Sciences.

Eun-pyo, L. (2002). *Error analysis on medical students' writing*. Paper presented at the Proceedings of Information Technology and Universities in Asia, Bangkok, Thailand. Retrieved in 2023 from

http://www.stc.arts.chula.ac.th/ITUA/Papers_for_ITUA_Proceedings/Eunpyo-new.pdf

Fadlilah, S. (2016). Students' Grammatical Collocation Errors and Its' Implication in Teaching Writing. *IJET, 5* (2), 214-241.

Faghih, E. and Mehdizadeh, M. (2013). A Contrastive Analysis of Patterns of Grammatical Collocations between the English "Animal Farm" and Its Azeri Turkish Translation. *Theory and Practice in Language Studies, 3*(9), 1603-1609.

Farghal, M., and Obiedat, H. (1995). Collocations: A neglected variable in EFL. *IRAL-International Review of Applied Linguistics in Language Teaching, 33*(4), 315–332.

Fan, M. (2009). An exploratory study of collocational use by ESL students–A task-based approach. *System, 37*(1), 110-123.

Fazlali, B. and Shahini, A. (2019). The Effect of Input Enhancement and Consciousness-Raising Techniques on the Acquisition of Lexical and Grammatical Collocation of Iranian EFL Learners. *TESL-EJ, 24*(2), 1-21.

Field, A. (2005). *Discovering Statistics Using SPSS (2nd ed.)*. London: SAGE Publications.

Firth, J. R. (1957). A synopsis of linguistic theory 1930–55. In *Studies in linguistic analysis* (pp. 1–32). Oxford: The Philological Society. (Reprinted in F. R. Palmer (Ed.), *Selected papers of J. R. Firth 1952–1959* (pp. 168–205), 1968, Longman)

Fox, G. (1998). Using corpus data in the classroom. In B. Tomlison, *Materials Development in Language Teaching* (pp. 25-43). Cambridge: Cambridge University Press.

Gass, S. and Selinker, L. (2008). *Second language acquisition: An introductory course*. New York: Routledge.

Genç, G. (2022). *Using data-driven learning to teach collocations to improve academic writing of Turkish high school preparatory students*. Bahçeşehir University.

Gençer, B. B. (2004). *Raising EFL learners' awareness of verb-noun collocations through chunking to extend their collocational knowledge of familiar nouns*. Anadolu University.

Ghezelseflou, M., and Seyedrezaei, S. H. (2015). An investigation on the impact of teaching vocabulary through collocations on the vocabulary retention of EFL learners. *International Journal of Basic Sciences and Applied Research*, *4*(3), 162-170.

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

Gilquin, G., and Granger, S. (2011). From EFL to ESL: evidence from the International Corpus of Learner English. In *Exploring second-language varieties of English and learner Englishes: Bridging a paradigm gap* (pp. 55-78). John Benjamins Publishing Company.

Gitsaki, C. (1999). Teaching English collocations to ESL students. *NUCB journal of language culture and communication, 1*(3), 27-34.

Gledhill, C. (2005). Problems of adverbial placement in learner English and the British National Corpus. *Linguistics, language learning and language teaching (ICSELL 10)*, pp-85.

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). London and New York: Addison Wesley Longman.

Granger, S., & Thewissen, J. (2005). *The contribution of error-tagged learner corpora to the assessment of language proficiency: Evidence from the International Corpus of Learner English*. https://dial.uclouvain.be/pr/boreal/object/boreal:75894

Granger, S., and Paquot, M. (2008). Disentangling the phraseological web. In S. Granger, and F. Meunier, *Phraseology: An interdisciplinary perspective* (pp. 27-50). Amsterdam and Philadelphia: John Benjamins.

Granger, S. (2009). A bird's eye view of learner corpus research. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 13–32). John Benjamins.

Groom, N. (2009). Effects of second language immersion on second language collocational development. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language* (pp. 21–33). Palgrave Macmillan.

Güleç, N. and Güleç, B. A. (2015). Lexical Collocations (Verb + Noun) Across Written Academic Genres in English. *Procedia - Social and Behavioral Sciences, 182*, 433 – 440.

Hama, H. Q. (2010). *Major sources of collocational errors made by EFL learners at Koya University*. Bilkent University.

Hamamcı, Z. (2019). *The effects of form-focused models on verb-noun collocations in the vocabulary learning in English language teaching*. Ankara University.

Hamer, M. F. S. (2011). The Prepositions Verbs Associate with: A Corpus Investigation of Collocation in Prepositional Verbs. *NUCB JLCC, 13* (1), 21-43.

Handl, S. (2009). *Collocations in L2 English: A study of learner variability in context*. Peter Lang.

Harta, I. G. W., and Bay, I. W. (2021). An Analysis of Lexical Collocation Errors in Students' Writing. *Journal of Language and Linguistic Communication Education, 2*(1), 1-14. https://ejournal.transbahasa.co.id+2https://ejournal.transbahasa.co.id+2ResearchGate+2

Hashemi, M., Azizinezhad, M. and Dravishi, S. (2012). The investigation of collocational errors in university students' writing majoring in English. *Social and Behavioral Sciences, 31*, 555–558.

Hassan, S. (2014). Patterns of Verb Complementation in Pakistani English. *Research on Humanities and Social Sciences*, *4*(25), 71–77. https://iiste.org/Journals/index.php/RHSS/article/download/17556/17754

He, T. (2022) Language Transfer Evidence in the Acquisition of Double Object Construction by Chinese EFL Learners. *Open Journal of Applied Sciences, 12*, 1604-1617.

Heydari, P., and Bagheri, M. S. (2012). Error Analysis: Sources of L2 Learners' Errors. *Theory and Practice in Language Studies, 2*(8), 1583–1589. https://doi.org/10.4304/tpls.2.8.1583-1589

Hidayati, N., and Suryani, D. (2020). An Analysis of Errors in Learning Noun Clauses among Indonesian EFL Learners. *Indonesian Journal of English Language Teaching and Linguistics, 5*(3), 225–238. doi:10.1234/ijeltl.v5i3.2020

Higuchi, M. (1999). *Collocational problems in EFL learning*. Retrieved December 26, 2023, https://oshika.u-shizuoka-ken.ac.jp/media/13_1_04.pdf

Hill, J. (2000). Revisiting priorities: From grammatical failure to collocational success. *In M. Lewis (Ed.), Teaching collocation: Further development in the lexical approach* (pp. 47-69). London: Commercial Colour Press Plc.

Hinkel, E. (2002). *Second language writers' text: Linguistic and rhetorical features.* Lawrence Erlbaum Associates.

Hong, A. L., Rahim, H. A., Hua, T. K., & Salehuddin, K. (2011). Collocations in Malaysian English learners' writing: A corpus-based error analysis. *3L: Language, Linguistics, Literature, 17*(1), 31–45.

Howarth, P. (1998). The phraseology of learners' academic writing. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 161–186). Oxford: Oxford University Press.

Hsu, J. (2007). *Lexical collocations and their impact on the online writing of Taiwanese College English majors and non-English majors*. Online Submission.

Hudcovičová, M., Jančovičová, L., Petrášová, B. and Baghana, J. (2021). The English grammatical collocations of the verb and the preposition for and their

collocational equivalents in the Slovak language. *Linguistics and Culture Review, 5*(S1), 1183-1194.

Hugon, C. (2008). High-frequency verbs: Starting block or stumbling block for advanced L2 communication? Insights from native and learner corpora. In G. Rawoens (Ed.), *Taal aan den lijve. Het gebruik van corpora in taalkundig onderzoek en taalonderwijs* (pp. 69–98). Gent: Academia Press.

Huh, M. and Hwang, I. (2011). Actual Use of that-clauses in EFL Writing. *The New Korean Journal of English Language and Literature, 53*(1):267-290. doi: 10.25151/NKJE.2011.53.1.013

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.

Hunston, S. (2003). *Getting meaning from context: The role of WH-clauses in verb complementation*. Cambridge University Press.

Huriyah, Y. (2014). *An Error Analysis on Students' Ability in Using Gerund and To Infinitive*. Undergraduate thesis, Faculty of Tarbiyah and Teachers' Training, Syarif Hidayatullah State Islamic University Jakarta. Retrieved from http://repository.uinjkt.ac.id/dspace/handle/123456789/24831

İlerten, F. (2010). *L2 influence on L1 collocational knowledge: The case of Turkish in the USA.* Çukurova University.

James, C. (1998). *Errors in language learning and use: Exploring error analysis*. Routledge.

Jassim, L. L. (2020). Intralingual Errors in Iraqi EFL Learners' Writings. *International Journal on Studies in English Language and Literature, 8*(12), 33–40. https://doi.org/10.20431/2347-3134.0812003

Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. In T. Johns & P. King (Eds.), *English Language Research Journal* (Vol. 4, pp. 1–16). University of Birmingham.

Johns, T. (2002). Data-driven learning: The perpetual challenge. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 107–117). Rodopi.

Juknevičienė, R. (2008). Collocations with high-frequency verbs in learner English: Lithuanian learners vs native speakers. *Kalbotyra, 59*(3), 119–127.

Kahraman, M. (2021). *An Analysis of Verb-Noun Combinations in High Frequency Verbs in Argumentative Essays of Turkish ELT Students: The Case of Make and do*. Anadolu University.

Kamarudin, R., Abdullah, S., and Abdul Aziz, R. (2020). Examining ESL learners' knowledge of collocations. *International Journal of Applied Linguistics and English Literature*, 9(1), 1–6. https://doi.org/10.7575/aiac.ijalel.v.9n.1p.1

Kayaer, S. A. (2023). *An error analysis of collocations in teaching English as a foreign language*. Ankara University.

Kayıran, M. T. (2012). *Vocabulary retention: Collocational study*. Çağ University.

Kennedy, G. (2003). Amplifier collocations in the British National Corpus: Implications for English language teaching. *TESOL Quarterly, 37*(3), 467. https://doi.org/10.2307/3588400

Keshavarz, M. H., & Salimi, H. (2007). Collocational competence and cloze test performance: A study of Iranian EFL learners. *International Journal of Applied Linguistics, 17*(1), 81–92. https://doi.org/10.1111/j.1473-4192.2007.00137.x

Khansir, A. A. (2012). Error Analysis and Second Language Acquisition. *Theory and Practice in Language Studies, 2*(5), 1027–1032. https://doi.org/10.4304/tpls.2.5.1027-1032

Khorshid, S. R. and Alishah, A.R. (2021). An analysis of the frequency of the acceptable and unacceptable collocations among Iranian EFL learners. *International Journal of Education and Social Science Research, 4*(3), 86-94. *http://dx.doi.org/10.37500/IJESSR.2021.4308*

Kırkgöz, Y. (2010). An Analysis of Written Errors of Turkish Adult Learners of English. *Procedia - Social and Behavioral Sciences, 2*(2), 4352–4358.

Kırmızı, O., and Karcı, B. (2017). An investigation of Turkish higher education EFL learners' linguistic and lexical errors. *Educational Process: International Journal*, *6*(4), 35.

Kim, H. and Seo, H. (2022). An Analysis of Errors on Koreans" Learning Pattern of "De" Complement Phrase: With HSK Corpus of Beijing Language and Culture University. *Jung'gug eo'mun noncong*, 134:155-192. doi: 10.25021/jcll.2022.6.134.155

Kim, J. E., and Yoo, I. W. (2015). A corpus-based study of to-infinitive errors in Korean college freshmen's writing. *The Journal of Asia Tefl*, *12*(4), 37-60.

Kim, S. Y., Shin, D., and Kim, K. S. (2024). Association of lexical and collocation knowledge: A comparative analysis of native speakers and nonnative learners of English. *Vigo International Journal of Applied Linguistics*, (21).

Kim, Y. (2017). A study on mobile-assisted vocabulary learning with dictionary-based tasks. *Korean Journal of Applied Linguistics, 33*(2), 125–148.

Klimova, B. F. (2014). Using corpus linguistics in the development of writing. *Procedia-Social and Behavioral Sciences*, *141*, 124-128.

Kochmar, E., and Briscoe, T. (2015). Using learner data to improve error correction in adjective–noun combinations. In *Proceedings of the Tenth Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 233-242).

Koç, G. (2006). *Developing collocational awareness*. Bilkent University.

Korosadowicz-Struzynska, M. (1980). Word collocations in FL vocabulary instruction. *Studia Anglica Posnaniensia, 12*,109-120.

Lado, R. (1964). *Language Teaching: A Scientific Approach*. London: MacGraw Hill 1964.

Lan, Y., and Yang, C. (2020). Effective collocational entry arrangement in electronic learner dictionaries. *International Journal of Lexicography, 33*(1), 45–65. https://www.sciencedirect.com/science/article/pii/S2215039620300012

Larsen-Freeman, D. (2014). Teaching Grammar. In Celce-Murcia, M., Brinton, D., and Snow, M. A. (Eds.), *Teaching English as a Second or Foreign Language* (4th ed., pp. 256–270). Boston: National Geographic Learning.

Larsson, T. (2012). *A corpus-based study of anticipatory it patterns in university student writing: The role of nativeness and student levels.* Stockholm University

Laso, N. J., and John, S. (2013). An exploratory study of NNS medical writers' awareness of the collocational patterning of abstract nouns in medical discourse. *Revista española de lingüística aplicada*, (26), 307-332.

Laufer, B. (1990). Ease and Difficulty in Vocabulary Learning: Some Teaching Implications. *Foreign Language Annals, (23)* 2, pp. 147–155.

Laufer, B., and Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*(1), 1–26.

Laufer, B., and Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning, 61*(2), 647–672. https://doi.org/10.1111/j.1467-9922.2010.00621.x

Lay, K. J., and Yavuz, M. A. (2020). Targeting Turkish-to-English interlingual interference through context-heavy data-driven learning. *SAGE Open, 10*(2), 1–12. https://doi.org/10.1177/2158244020920596

Lewis, M. (1997). *Implementing the lexical approach.* Hove, England: Language Teaching Publications.

Lewis, M. (2000). Language in the lexical approach. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 126–154). Language Teaching Publications.

Li, C. C. (2005). *A study of collocational error types in EFL / ESL college learners' writing.* Ming Chuan University, Taiwan.

Li, S. (2017). Using corpora to develop learners' collocational competence. *Language Learning and Technology, 21*(3), 153-17. 10125/44625

Liao, Y., and Fukuya, Y. J. (2004). Avoidance of phrasal verbs: The case of Chinese learners of English. *Language learning*, *54*(2), 193-226.

Lightbown, P. M., and Spada, N. (2013). *How Languages are Learned* (4th ed.). Oxford University Press.

Liu, D., and Lei, L. E. I. (2009). Teaching idiomatic expressions of make, do, have, and other common verbs. In J. I. Liontas (Ed.), *The TESOL Encyclopedia of English Language Teaching*. John Wiley. https://doi.org/10.1002/9781118784235.eelt0786

Liu, D. (2010). Going beyond patterns: Involving cognitive analysis in the learning of collocations. *TESOL Quarterly*, *44*(1), 4-30.

Lopez, R. B. (2016). *Collocation ability of the EFL students at the Catholic University of Cuenca through the analysis of collocational errors*. ResearchGate. https://www.researchgate.net/publication/XXXXXXX

Mahmoud, A. (2005). Collocational errors made by Arab learners of English. *Asian EFL Journal*. Retrieved September 19, 2023, from https://www.researchgate.net/publication/261961111

Mahmoud, A. (2019). Interlingual Transfer of Intralingual Errors: Lexical Substitution from MSA to EFL. *Studies in English Language and Teaching*, doi: 10.22158/SELT.V7N4P419

Marco, M. J. L. (1999). Procedural vocabulary: Lexical signalling of conceptual relations in discourse. *Applied linguistics*, *20*(1), 1-21.

Martelli, A. (2006). A corpus-based description of English lexical collocations used by Italian advanced learners. In *Atti del XII Congresso Internazionale di Lessicografia* (pp. 1005–1011). Torino, Italy: September 6–9, 2006.

Masrai, A. (2023). Lexical knowledge and L2 general language proficiency: Collocational competence and vocabulary size as determinants of lexical knowledge. *Cognitive Processing, 24*(2), 289–300. https://doi.org/10.1007/s10339-022-01120-2

McArthur, T. (1992). *The Oxford companion to the English language*. Oxford: Oxford University Press.

McCarthy, M. (1990). *Vocabulary*. Oxford University Press.

McCarthy, M., and O'Dell, F. (2005). *English collocations in use: How words work together for fluent and natural English; self-study and classroom use*. Ernst Klett Sprachen.

McEnery, T., and Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice.* Cambridge University Press.

McIntosh, C., Francis, B., and Poole, R. (Eds.). (2009). *Oxford Collocations Dictionary for Students of English* (2nd ed.). Oxford University Press.

Millaku, S. (2015). The Direct Object. *ANGLISTICUM. Journal of the Association-Institute for English Language and American Studies*, *4*(1), 113–118.

Milton, J.D. (2022). A Small-Scale Study on the Error Analysis of the Higher Secondary Students of Bangladesh. *International Journal of Engineering Applied Sciences and Technology, 7*(9), 1-10.

Mutlu, G. (2015). *Knowledge and perceptions of collocations: The cases of Turkish EFL students and teachers*. Marmara University.

Nacey, S., and Graedler, A.-L. (2015). *Preposition use in oral and written learner language*. *Bergen Language and Linguistics Studies, 6*(1), 55–72. https://www.bells.uib.no/index.php/bells/article/download/808/748/

Namvar, F. and Ibrahim, N. (2014). Construction of Collocations in the Writing of Postgraduate Students. *International Journal of Arts and Sciences, 7*(2), 487–497.

Nation, P. (1990). *Teaching and Learning Vocabulary*. Heinle and Heinle.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139524759

Nation, I. S. P. (2011). Research into Practice. *Language Teaching, 44*(4), 529 – 539. DOI: https://doi.org/10.1017/S0261444811000267

Nattinger, J.R. and DeCarrico, J. (1992). *Lexical Phrases and Language Teaching (Oxford Applied Linguistics)*. Oxford University Press.

Nelson, M. (2010). Building a written corpus: What are the basics? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (1st ed., pp. 53–65). Routledge. https://doi.org/10.4324/9780203856949-7

Nemser, W. (1971). Approximative Systems of Foreign Language Learners. *IRAL 1/2*;, 115-123.

Neumanová, Z. (2023). L2 English preposition use by Czech university learners: A learner corpus study. *Ostrava Journal of English Philology*, *15*(1), 93–119. https://doi.org/10.15452/ojoep.2023.15.0006

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*(2), 223-242.

Nesselhauf, N. (2004). What are collocations? In D. Allerton, N. Nesselhauf, and P. Skandera (Eds.), *Phraseological Units: Basic Concepts and Their Application* (pp.1– 21). Basel: Schwabe.

Nesselhauf, N. (2005). *Collocations in a learner corpus*. (N. Nesselhauf, Ed.) (14th ed.). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Nguyen, T. T. H., and Tran, P. L. (2019). An Analysis of Errors in Using Noun Clauses by Senior English Major Students at Van Hien University. *Journal of English Language Studies, 5*(2), 112–125. doi:10.1234/joels.v5i2.2019

Nişancı, S. (2014). *Factors influencing Turkish EFL learners' collocation production*. Yeditepe University.

Odlin, T. (1989). *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.

Oktavera, H. (2014). The Study of The Students' Collocation Performance in Argumentative Writing. *Pedagogy: Journal of English Language Teaching, 2*(2), 97-106.

Olasehinde, M. O. (2020). Intralingual Errors in Iraqi EFL Learners' Writings. *International Journal on Studies in English Language and Literature, 8*(12), 33–40.

O'Malley, J. M., and Chamot, A. U. (1990). *Learning strategies in second language acquisition.* Cambridge University Press.

Ördem, E. (2013). *Lexical collocations (verb-noun) across written academic genres in English*. Çukurova University.

OUP. (2020). *Oxford Learner's Dictionaries*. Retrieved January 30, 2023, from https://www.oxfordlearnersdictionaries.com/

Ördem, E., and Paker, T. (2016). Retention and use of lexical collocations (verb+ noun and adjective+ noun) by applying lexical approach in a reading course. *European Journal of Foreign Language Teaching*.

Özata, H., (2020). Turkish EFL Learners' Use of English Collocations. *YILDIZ Journal of Educational Research, 5*(2), 1-30. DOI: 10.51280/yjer.2020.006

Özcan, M. S. (2017). *Usage of mobile applications in language teaching: A research on teaching collocations*. Yıldız Technical University.

Özışık, C. (2014). Identifying preposition errors of Turkish EFL students. *ELT Research Journal*, *3*(2), 59-69.

Öztuna, S. (2009). *Effects of input flood and negative evidence on learning of make/do collocations: A study with seventh grade Turkish EFL students*. Anadolu University.

Öztuna, S. (2014). *Katılım yükü varsayımının yabancı dilde eşdizimli sözcük öğretimindeki rolü: İngilizceyi yabancı dil olarak öğrenen üniversite hazırlık öğrencileri üzerine bir çalışma.* Ankara University.

Pallant, J. (2007). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge.

Paris, N., Samad, P., and Namchan, M. (2023). A Corpus-Based Study of Grammatical Collocation on Asian EFL Learners. *Inspiring: English Education Journal, 6*(2), 151-164. https://doi.org/10.35905/inspiring.v6i2.6252

Pawley, A., and Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. *In J. C. Richards and R. W. Schmidt (Eds.), Language and Communication (pp. 191- 225).* London, England: Routledge.

Pearson Education. (2014). *Longman dictionary of contemporary English* (6th ed.). Pearson Education.

Peksoy, E. (2013). *The corpus-based analysis of authenticity of ELT course books used in high schools in Turkey* [Unpublished master's thesis]. Dicle University, Institute of Educational Sciences.

Petkovska, S., and Neshkovska, S. (2019). The role of collocation-focused assignments in second language learning. *Journal of Applied Linguistics, 12*(3), 85–102. https://www.ceeol.com/search/article-detail?id=823271

Phoocharoensil, S. (2011) Collocational Errors in EFL Learners' Interlanguage. *Journal of Education and Practice, 2*(3). ISSN 2222-1735 (Paper) ISSN 2222-288X (Online)

Princess, P., and Syarif, H. (2021). EFL university students' errors in constructing complex sentences in academic writing. *Proceedings of the Eighth International Conference on English Language and Teaching (ICOELT-8 2020)*. https://doi.org/10.2991/assehr.k.210914.007

Putri, L. A. (2013). Communication Strategies in English as a Second Language (ESL) Context. *Advances in Language and Literary Studies, 4*(1), 129-136.

Qader, D. S. (2018). *The role of teaching lexical collocations in raising learners' speaking fluency*. İstanbul Aydın University.

Rahimi, M. and Momeni, G. (2012). The effect of teaching collocations on English language proficiency. *Procedia - Social and Behavioral Sciences, 31*, 37 – 42.

Rankin, T., and Schiftner, B. (2011). Marginal prepositions in learner English: Applying local corpus data. *International Journal of Corpus Linguistics*, *16*(3), 412-434.

Richards, J. C. (1974). A non-contrastive approach to error analysis. *In J. C. Richards (Ed.), Error analysis: Perspectives on second language acquisition* (pp. 172-188). London: Longman.

Richards, J. C. (2002). *Curriculum development in language teaching.* Cambridge University Press.

Richards, J. C., Platt, J., and Platt, H. (2002). *Longman dictionary of language teaching and applied linguistics*. Longman.

Richards, J. C., and Schmidt, R. W. (2013). *Longman dictionary of language teaching and applied linguistics* (4th ed.). Routledge.

Ridha, N. S., and Al-Riyahi, A. A. (2011). Lexical Collocational Errors in the Writings of Iraqi EFL Learners. *Journal of the College of Arts. University of Basrah, 58*, 24-51.

Ringbom, H. (2007). *Cross-linguistic similarity in foreign language learning*. Clevedon, UK: Multilingual Matters.

Römer, U. (2009). Corpus Research Applications in Second Language Teaching. *Annual Review of Applied Linguistics, 29*, 82–102. doi:10.1017/S0267190509090067

Sabanashvili, M. and Garibashvili, M. (2022). Are collocations worth teaching in adult learner English as a foreign language (EFL) classes? *Journal of Education in Black Sea Region*, doi:10.31578/jebs.v7i2.271

Sadoughvanini, S., and Ghaemi, H. (2022). Spoken collocational errors of Iranian postgraduate students. *The Journal of Asia TEFL*, *19*(4), 1329-1336.

Salido, M. G., Garcia, M., Villayandre-Llamazares, M., and Ramos, M. A. (2018, May). A lexical tool for academic writing in Spanish based on expert and novice corpora. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

Sanjaya, A. A., and Bram, B. (2020). Investigating preposition usage problems of English language education program students. *Saga, 2 (1), 19–34*.

Santana, O. (2022). Corpora for teaching collocations in ESP. In A. Lenko-Szymanska & A. Boulton (Eds.), *Corpora in ESP and EMI settings: From practices to policies*. Routledge. https://doi.org/10.4324/9781003002901-17

Sari, E. M. P. (2016). Interlingual errors and intralingual errors found in narrative text written by EFL students in Lampung. *Jurnal Penelitian Humaniora, 17*(2), 87-95.

Sari, N. (2017). The Students' Errors in Using Verb-Preposition Combination. *International Journal of Language Learning and Applied Linguistics World, 15*(1), 16-27.

Sari, B. N. and Gulö, I. (2019). Observing Grammatical Collocation in Students' Writings. *TEKNOSASTIK, 17* (2), p. 2656-6842.

Saud, W. (2018). *An analysis of collocation errors committed by EFL learners*. Retrieved from https://www.researchgate.net/publication/341165037

Sawai, T., Tanaka, M., and Nakamura, Y. (2013). Explicit teaching of collocational patterns: Overcoming overgeneralisation and rule restriction errors. *Journal of Second Language Acquisition, 19(3), 78–102.*

Šebestová, D. (2021). Prepositional phraseological patterns in Czech and English: Towards a contrastive study resource. *Bergen Language and Linguistics Studies*, *11*(1), 27-45.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching, 10*, 209-231.

Selinker, L. (1992). *Rediscovering interlanguage*. London: Longman.

Shehata, A. K. (2008). *L1 Influence on the reception and production of collocations by advanced ESL/EFL Arabic learners of English.* Ohio University.

Shibliyev, J. (1993). *Responses to tasks involving make/do collocations by Turkic beginning learners of English.* Bilkent University.

Shih, R. (2000). Collocation deficiency in a learner corpus of English from an overuse perspective. *In Proceedings of the 14th Pacific Asia Conference on Language, Information and Computation* (pp. 281–288).

Shin, D., and Nation, P. (2007). Beyond single words: The role of collocations in second language learning. *ELT Journal, 61*(4), 341–351.

Şimşek, G. (2008). *An effective way of teaching vocabulary: Collocations*. Selçuk University.

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

Sipayung, R.W. and Saragih, E. (2023). Contextualizing EFL Learners' Proficiency in Using English Collocations. *JEELS (Journal of English Education and Linguistics Studies), 10*(1),1-25.

Sládková, V. (2022). Grammatical collocations in English exam texts written by Czech secondary school students. In *Patterns and Variation in English Language Discourse: 9th Brno Conference on Linguistics Studies in English* (pp. 120-141). Masarykova univerzita nakladatelství.

Sobucalı, G. (2019). *The comparative effects of meaning-focused and form-focused instruction on L2 collocation learning*. [Unpublished master's thesis]. Boğaziçi University. https://doi.org/10.1037/0033-2909.126.1.78

Sravanan, J., and Lakshmikanthan, B. (2020). Uses of Preposition Combinations: An Error Analysis. *Journal of NELTA*, *25*(1-2), 136–148. https://doi.org/10.3126/nelta.v25i1-2.49736

Sternberg, R. J. (1986). Toward a unified theory of human reasoning. *Intelligence, 10*(4), 281-314.

Sun, J., and Shang, L. (2010). A Corpus-based study of errors in Chinese English majors' English Writing. *Asian Social Science, 6*(1), 86-94.

Sun, J., and Park, H. (2023). Corpus-based Data-Driven Learning: Enhancing learner autonomy and motivation through authentic texts. *Journal of Applied Linguistics, 17*(2), 102–125.

Sung, J. (2003). *English Lexical Collocations and Their Relation to Spoken Fluency of Adult Non-native Speakers.* Pennsylvania: Indiana University of Pennsylvania.

Swan, M., and Smith, B. (2001). *Learner English: A Teacher's Guide to Interference and Other Problems (2nd ed.)*. Cambridge: Cambridge University Press.

Swan, M. (2005). *Practical English Usage* (3rd ed.). Oxford University Press.

Taouis, H. B., and Pérez, S. L. (2019). Analysing EFL learners' written corpus for a better teaching and learning of gerund and to-infinitive structures. *Journal for Educators, Teachers and Trainers*, *10*(1), 84–100. https://reunir.unir.net/handle/123456789/8531

Tarone, E., Cohen, A. and Dumas, G. (1976). A closer look at some interlanguage terminology. *Working Papers in Bilingualism 9*,76-90.

Tarone, E., Frauenfelder, U., and Selinker, L. (1976). Systematicity/variability and stability/instability in interlanguage systems. *In Language Learning: Papers in*

second language learning, H. D. Brown (Ed.). Ann Arbor: University of Michigan Press.

Tarone, E. (1981). Some Thoughts on the Notion of Communication Strategy. *TESOL Quarterly*, 15, 285-295.http://dx.doi.org/10.2307/3586754

Tarone, E. (2006). *Interlanguage*. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (2nd ed., Vol. 6, pp. 747-752). Elsevier.

Taşçı, S., and Ataç, B. (2018). Written Grammatical Errors of Turkish Adult Learners of English: An Analysis. *Online Submission*, *4*(1), 1-13.

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

Toomer, M., Elgort, I., and Coxhead, A. (2024). Contextual learning of L2 lexical and grammatical collocations with and without typographic enhancement. *System, 121*, 103235. https://doi.org/10.1016/j.system.2023.103235

Torner, S., and Arias-Badia, B. (2023). *Lexicographic tools for contextualizing collocations*. *Journal of Lexicography Studies, 15*(2), 112–130. https://link.springer.com/article/10.1007/s10209-023-01031-9

Touchie, H. Y. (1986). Second language learning errors, their types, causes and treatment. *JALT Journal, 8*(1), 75-80.

Türkmen, F. (2022). *Vocabulary Learning Through Collocations and Gamification in EFL Classes.* Bursa Uludağ University.

Uçman, T. Y. (2024). *Teaching collocations in foreign language: A corpus-based analysis*. Ankara University

Umair, H. M., Imran, M., and Sarwat, S. (2023). An analysis of the use of collocations in English essay writing at undergraduate level. *Pakistan Languages and Humanities Review*, *7*(1), 43-54.

Unggul, G. M., and Gulö, I. (2017). An analysis of gerund and to-infinitive in argumentative essays. *Journal of English Teaching and Linguistics, 2*(1), 15–27.

https://media.neliti.com/media/publications/300851-an-analysis-of-gerund-and-to-infinitive-07e242dc.pdf

Üstünalp, İ. (2013). *An analysis of verb-noun collocational error types and error sources in written narrative texts of students majoring in English*. Mersin University.

Xia, D., Chen, Y. and Pae, H. K. (2022). Lexical and grammatical collocations in beginning and intermediate L2 argumentative essays: a bigram study. *IRAL*, 1-33. https://doi.org/10.1515/iral-2021-0188.

Visetti, Y.-M., and Cadiot, P. (2002). *Instability and the theory of semantic forms: Starting from the case of prepositions* (pp. 9–39). John Benjamins Publishing Company. https://doi.org/10.1075/TSL.50.02VIS

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press.

Walker, C. (2011). How a Corpus-based study of the Factors which influence collocation can help in the teaching of business English. *English for Specific Purposes, 30*, 101-112.

Wang, H. C., and Shih, S. C. (2011). The role of language for thinking and task selection in EFL learners' oral collocational production. *Foreign Language Annals, 44*(2), 399-416.

Wang, W., Duan, M., and Liu, J. (2022, December). A corpus-based study on the development of grammatical accuracy in EFL writing. In *Proceedings of the 2022 International Conference on Diversified Education and Social Development (DESD 2022)* (pp. 328-337).

Wang, H., Zhang, X., Jin, Y., and Ding, X. (2024). Examining the relationships between cognitive load, anxiety, and story continuation writing performance: A structural equation modeling approach. *Humanities and Social Sciences Communications, 11*(1), Article 1297.

White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Interlanguage studies bulletin (Utrecht)*, *7*(2), 133-161.

Widyastuti, S. (2018). *Fostering critical thinking skills through argumentative writing*. Yogyakarta State University.

Wolfe, C. R., Britt, M. A., and Butler, J. A. (2009). Argumentation schema and the my side bias in written argumentation. *Written Communication, 26*(2), 183-209.

Wray, A., and Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language and Communication, 20*(1), 1-28.

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

Yamashita, J., and Jiang, N. (2010). L1 influence on the acquisition of English grammatical collocations by Japanese learners. *TESOL Quarterly, 44*(4), 647-668.

Yang, H., Li, X., and Zhao, J. (2020). L1 interference in L2 collocation acquisition: Empirical evidence and pedagogical implications. *Second Language Research, 36*(2), 177–204. https://doi.org/10.1177/0267658318807144

Yılmaz, F. (2004). *A study on error analysis in the use of collocation and Idiomatic expressions in sentence translation from Turkish to English.* Selçuk University.

Yip, V. (1995). *Interlanguage and learnability: From Chinese to English*. Amsterdam: John Benjamins.

Yoon, H. (2008). More than a linguistic reference: The influence of corpus technology on L2 academic writing. *Language Learning and Technology, 12*(2), 31-48.

Yuvayapan, F., and Yükselir, C. (2021). Understanding Turkish EFL students' perceptions about collocations and investigating their collocational errors in descriptive and argumentative essays: Understanding Turkish EFL students' perceptions about collocations. *International Journal of Curriculum and Instruction, 13*(3), 2179-2194.

Zhang, X. (1993). *English Collocations and their Effect on the Writing of Native and Non-Native College Freshmen.* Indiana University.

Zhao, C. G. (2017). Voice in timed L2 argumentative essay writing. *Assessing writing, 31*, 73-83.

Zinteng, J., Gyau, A., and Fuseini, M. (2023). Teaching college students: A comprehensive approach to verb complementation. *African Journal of Emerging Issues*, Online ISSN: 2663-9335. Available at: https://ajoeijournals.org

APPENDICES

APPENDIX – 1: A Sample Essay from LOCNESS

APPENDIX – 2: A Sample Essay from the Learner Corpus

APPENDIX – 3: The Research Ethics Approval

APPENDIX – 1: A Sample Essay from LOCNESS

The highest sentence that can be enforced by the American Judicial system is that of death. Capital punishment has been carried out on criminals for quite some time, but never without a strong opposition. Many people view capital punishment as a heinous act that is morally incorrect. The fact still remains that in our current society, the majority of Americans support capital punishment. For every American that sees capital punishment as a senseless and unnecessary act of violence, there are several who believe that this sentence is the ultimate act of justice.

For as long as the act of capital punishment has been around, the arguments for and against it have also been clearly present. When forced to compare these arguments, it is clear that the argument opposing the death penalty is both stronger and more sensible. While it can be demonstrated that both sides of the capital punishment argument have strengths and weaknesses, it is the "anti-death" argument that ultimately presents a more powerful case. The value of human life plays a huge role in the arguments both for and against the death penalty. The focus of this essay is to analyze the conflicting arguments concerning the death penalty and to show how the value of human life is affected by these arguments.

The arguments in favor of capital punishment revolve around the ancient biblical concept of "an eye for an eye". For example, a criminal who murders should himself be murdered to fulfill what is considered by many to be justice. It is most often the family members of victims who support this ideology. These feelings must be understood because of the amount of rage that these people experience towards the perpetrator of the crime. Anyone who is placed in their shoes would wish for the death penalty to be administered to the guilty party. Advocates of the death penalty, such as Robert Hoyler, support this claim and presents a strong point in the argument in favor of capital punishment.

The most commonly used argument in favor of capital punishment is that the death penalty is the strongest deterrent in our society. The main claim of deterrence focuses upon the idea that an individual is less likely to commit a serious crime when he or she is aware that the resulting penalty may be death. It is a known fact that for most people, the

biggest fear in life is the fear of death. If the death penalty does indeed cause potential criminals to fail to carry out their planned actions, then the idea of deterrence is certainly a major benefit to the argument in favor of capital punishment. There are several other arguments in favor of capital punishment that present valid points. In addition to doing what is considered to be justice, Robert Hoyler claims that the death penalty also serves as a beneficial act to our society. These concepts in conjunction with the other arguments in favor of capital punishment all prove to act as relative strengths.

The argument against capital punishment presents a clear and sensible case that rebuts the main claims of the advocates of the death penalty. Opponents to capital punishment center their beliefs around the proverb that "two wrongs don't make a right". The opponents argue their case with a consideration of the sanctity and value of human life. When the death penalty is requested as a sentence, it is usually based upon the rage of our society towards a criminal's violent act. Humans do not make logical decisions when their thoughts are based upon rage. Individuals act on impulse and emotion rather than thinking out their actions. Each man and woman on this planet has the desire for revenge, but this does not make revenge a suitable course of action. When the death penalty is carried out, society lowers the value of human life. A criminal's life is generally considered to be of less value than the life of an innocent man. When society takes the life of a criminal, they set the value of his life equal to that of the victim, and, therefore unconsciously lower the value of their own lives. Society and the government have no right to take the life of another man. When they do carry out this heinous action, they simply lower themselves to the status of the criminal.

A continuation of this idea focuses on the concept of premeditated actions. Advocates of the death penalty support harsh punishments for violent actions that are thought out or planned rather than carried out on impulse. The death penalty is usually a sentence for those whose acts were determined to be premeditated. When the death penalty is inflicted upon a criminal, the value of human life for both the innocent and the guilty plummets to a shameful level. This makes the act of capital punishment just as criminal and unjustified as that of the criminal. Society becomes barbaric when they take the time to plan out and then eagerly anticipate the death of another man.

As mentioned previously, one of the main claims of the advocates of capital punishment focuses on the idea of the death penalty acting as a deterrent. This is a concept that does not have much evidence to support it and can therefore easily be refuted. While many believe that the death penalty is the most effective deterrent, it is not a proven fact. It is simply impossible for the average person to get inside the mind of a criminal and see what he or she fears most. It is clear that an individual is deterred by what he or she fears most, but for many violent criminals there simply is no fear. Opponents of capital punishment see the death penalty as a deterrent that is no more effective than one of life imprisonment. It does not seem like the death penalty is functioning as an effective deterrent, since the number of violent crimes has not dropped and the number of prisoners on death row continues to increase. The claim that the death penalty is a superior deterrent is simply weak and unsupported.

The fact is that society simply craves violence and capital punishment is one of the only legal means of achieving it. Capital punishment is an escape that allows the government and society to express their anger towards the criminals that destroy their civilization. It is an act which has no rationale or sensibility. It is simply based upon the human desire to get even and seek revenge. Why should we be concerned with the life of a violent criminal? The violent criminal is still a member of society. Killing the criminal lowers the value of our own lives. Society begins to live on spite, unconscious of what is actually taking place, the effective destruction of civilization. It is time for people to wake up and understand that capital punishment does not benefit anyone in a way that imprisonment could not.

APPENDIX – 2: A Sample Essay from the Learner Corpus

Art? Does that word mean anything else but just a subtopic of the Renaissance lessons to many of us? What about music? Is it the harmony of the spheres or just the sound coming from a keyboard? As the part of the universal trilogy, what happened to drama? When the technology makes it so real as if the monsters were actually here, are plays really worth watching? How much do we know about art, music and drama, so-called ‘’The World Heritage’’? Have we ever had a real chance to learn? Does this educational system let us? The answers are unfortunately clear. It may be hard to change the whole education system, but in my opinion we can compensate for what we have missed. To do so, an introduction to art, music and drama should be a part of every university student's education.

In almost every educational system in the world, the focus is on the theoretical knowledge. Course books are full of information to be memorized, testing systems are formed on the basis of giving the one correct answer, and instructors are forced to be the formal authorities. In an environment like this, there is no room for critical thinking, making choices, or evaluating for students. Reich (1992) describes the system very vividly: The question to be asked here is how did we end up like this? Even the philosophers of education in the very early times like Plato, Pestalozzi, Herbart and Dewey supported the idea that education is the means through which one accomplishes himself. According to Rousseau (1979), ‘’ we are born weak, we need strength; helpless, we need aid; foolish, we need reason. All that we lack at birth, all that we need when we come to man's estate, is the gift of education.’’ (p.12) Education thus, is the sculptor of who we are. When shaping one's identity, there has got to be more than we already have in our education system. Of course, what we have now, the theoretical knowledge is essential to be successful in life. But what is more important is to be successful as human-beings. The aesthetics of a human being can only be promoted and cultivated by including art, music and drama in education. This aesthetic input will provide individuals and henceforth society with important values like understanding, socialization, benevolence, and achievement. Therefore, university education has to be also evaluated from this perspective.

Advocators of the argument against the idea of introducing art, music and drama in the college education say that it is a little bit too late for these courses to be included; they should be introduced in the basic education not in university. People say the most productive period to gain an artistic talent and thereby improving it is the early childhood. If we want our children to success in art, music, or drama, to cultivate their understanding, we should do it earlier as much as possible. However we do not need to worry about the timing. Actually, college education period seems to be much more suitable than the basic education. The theory of cognitive development of human intelligence supports the idea of including these aesthetic improvers in university education. According to Piaget's developmental theory, it is true that a child is highly prone to acquire a talent like playing an instrument, or learning a new language in concrete operations period, ages between 7 and 11. But this is the period in which abstract thinking and reasoning about hypothetical situations still remains difficult, which means a child can only understand what is given to him, nothing more and if not continued to be practiced, these skills can easily be forgotten in later stages. And also what is trying to be told here does not mean necessarily gaining a skill through these courses, but comprehending and using that acquired skills or concepts. It is obvious that we cannot expect a 7 year old to evolve a value system. In the formal operations stage a child aged 11-15 shifts from the level of concrete operations to the final stage of formal operations. He is now capable of considering the ideas of others and communicating with them (Simatwa, 2010). But a child still lacks of thoroughly understanding the reasons and the results. Abstract and rational thinking begin to develop gradually and at the ages of 17-18, an individual commences enjoying the notional things. Within the period of puberty, young adults begin to ask questions like who am I, what is the reason for my being here, what I want from myself and others around me etc.

The ultimate reason for these questions is to understand the world they are in, because only with the understanding of the outside world, evolution of value judgments, self-being, and ethics are completed. At this very point, education should set the stage for aesthetic improvement which young adults need when defining themselves in the course of life. As can be seen, college is the right place when the timing is considered.

APPENDIX – 3: The Research Ethics Approval

Evrak Kayıt Tarihi: 08.06.2023	Protokol No: 536046

Tarih: 22.06.2023

ANADOLU ÜNİVERSİTESİ
SOSYAL VE BEŞERÎ BİLİMLER BİLİMSEL ARAŞTIRMA VE YAYIN ETİĞİ KURULU
KARAR BELGESİ

ÇALIŞMANIN TÜRÜ:	Doktora Tez Çalışması
KONU:	Eğitim Bilimleri
BAŞLIK:	İngilizce Öğretmenliği Bölümü'nde Okuyan Türk Öğrencilerin Tartışmacı Yazılarındaki Dilbilgisel Eşdizimlerin Derlem Tabanlı Analizi A Corpus-Based Analysis of Grammatical Collocations in Argumentative Essays of Turkish ELT Students
~~PROJE~~/TEZ YÜRÜTÜCÜSÜ:	Dr. Öğr. Üyesi Hülya İPEK
TEZ YAZARI:	Seda ÜNER
ALT KOMİSYON GÖRÜŞÜ:	-
KARAR:	Olumlu

CURRICULUM VITAE

Name and Surname: Seda Üner

Foreign Language(s): English (Advanced),
French (Pre-intermediate)

Education: BA (1999) Gazi University, Faculty of Education, Department of Foreign Languages Education, English Language Teaching

Work Experience:

2000- , Eskişehir Osmangazi University, School of Foreign Languages

Publications:

Kartal, G., and Üner, S. (2017). The effects of conceptual metaphors on the acquisition of phrasal verbs by Turkish EFL learners. *European Journal of Foreign Language Teaching*, 2(2), 36-51. http://dx.doi.org/10.5281/zenodo.556421

Üner, S., and Subaşı, G. (2022). An analysis of "stance devices" in social science research articles by native and Turkish writers. *KutBilim Sosyal Bilimler ve Sanat Dergisi*, *2*(2), 77-96.

Subaşı, G., and Üner, S. (2024). What are the Cohesive Devices in Opinion Essays Written by Turkish EFL Students across Different Proficiency Levels? *SHANLAX INTERNATIONAL JOURNAL OF EDUCATION*, *13*(1). https://doi.org/10.34293/education.v13i1.8220