

KARADENİZ TECHNICAL UNIVERSITY * THE INSTITUTE OF SOCIAL SCIENCES

DEPARTMENT OF WESTERN LANGUAGES AND LITERATURE

APPLIED LINGUISTICS MASTER'S PROGRAM

**A CORPUS ANALYSIS OF LEARNER ENGLISH FROM DIFFERENT BACKGROUNDS:
INVESTIGATING VERB-NOUN PATTERNS THROUGH THE PATTERN DICTIONARY
OF ENGLISH VERBS**

MASTER'S THESIS

Buse UZUNER

JULY - 2023

TRABZON

KARADENİZ TECHNICAL UNIVERSITY * THE INSTITUTE OF SOCIAL SCIENCES

DEPARTMENT OF WESTERN LANGUAGES AND LITERATURE

APPLIED LINGUISTICS MASTER'S PROGRAM

**A CORPUS ANALYSIS OF LEARNER ENGLISH FROM DIFFERENT BACKGROUNDS:
INVESTIGATING VERB-NOUN PATTERNS THROUGH THE PATTERN DICTIONARY
OF ENGLISH VERBS**

MASTER'S THESIS

Buse UZUNER

Supervisor: Assoc. Prof. Dr. Ali Şükrü ÖZBAY

JULY, 2023

TRABZON

APPROVAL

Upon the submission of the dissertation, **Buse UZUNER** has defended the study “**A Corpus Analysis of Learner English from Different Backgrounds: Investigating Verb-Noun Patterns through the Pattern Dictionary of English Verbs**” in partial fulfilment of the requirements for the degree of Master of Arts in Applied Linguistics at Karadeniz Technical University, and the study has been found fully adequate in scope and quality as a thesis by **unanimous / majority** vote on **03/07/2023**.

Committee Member		Decision		Signature
Title - Name and SURNAME	Mission	Accept	Refuse	
Assoc. Prof. Dr. Ali Şükrü ÖZBAY	Supervisor	☐	☐	
Assoc. Prof. Dr. Ceyhun YÜKSELİR	Member	☐	☐	
Asst. Prof. Dr. Hasan SAĞLAMEL	Member	☐	☐	

Approval of the Graduate School of Social Sciences.

Prof. Dr. Tülay İLHAN NAS
Director

DECLARATION OF ORIGINALITY

I, **Buse UZUNER**, hereby confirm and certify that;

- I am the sole author of this work and I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution,
- this work contains no material that has been submitted or accepted for a degree or diploma in any other university or institution,
- all data and findings in the work have not been falsified or embellished,
- this is a true copy of the work approved by my advisor and thesis committee at Karadeniz Technical University, including final revisions required by them.
- I understand that my work may be electronically checked for plagiarism by the use of plagiarism detection software and stored on a third party's server for eventual future comparison,
- I take full responsibility in case of non-compliance with the above statements in respect of this work.

Signature.....

03/07/2023

ACKNOWLEDGEMENT

First and foremost, I would like to thank my supervisor, Assoc. Dr. Ali Şükrü ÖZBAY, for his guidance, reassurance, and encouragement throughout the thesis period. He has always been there to help me in any case and encourage me to become a good researcher. I would like to thank Assist. Prof. Hasan SAĞLAMEL for his enormous contributions to my writing and academic skills until the prep year and being a jury member of my thesis. I want to thank Assoc. Dr. Ceyhun YÜKSELİR for accepting to be a jury member of my thesis and his valuable comments about the design and content of my thesis. I also want to thank Assist Prof. Dr. Muhammed BAYDERE for being one of my jury members and sharing his invaluable comments and questions about the study. I want to share my deepest gratefulness to Prof. Paola CATENACCIO for providing me the opportunity to work on corpus linguistics under her supervision during my days in University of Milan. I also want to thank Res. Asst. Tuncer AYDEMİR for his friendship and valuable support to my academic works.

This work was supported by Office of Scientific Research Projects of Karadeniz Technical University. Project number: 10667

July, 2023

Buse UZUNER

TABLE OF CONTENTS

ACKNOWLEDGEMENT	IV
TABLE OF CONTENTS.....	V
ÖZET.....	VII
ABSTRACT	VIII
LIST OF TABLES	IX
LIST OF FIGURES	X
LIST OF ABBREVIATIONS.....	XI
INTRODUCTION.....	1-2

CHAPTER ONE

1. FRAMEWORK OF THE STUDY.....	3-7
1.1. Introduction.....	3
1.2. Purpose of the Study	4
1.3. Significance of the Study	4
1.4. Statement of the Problem	5
1.5. Research Questions	6
1.6. Organisation of the Thesis	6

CHAPTER TWO

2. LITERATURE REVIEW.....	8-18
2.1. Introduction.....	8
2.2. The Notion of Pattern.....	8
2.3. Corpus Pattern Analysis (CPA)	10
2.3.1. Collocation and CPA	11
2.3.2. Colligation and CPA	12
2.3.3. Semantic Preference and CPA	12
2.4. Pattern Dictionary of English Verbs (PDEV)	12
2.4.1. Identifying Verb Patterns	13
2.4.2. The Disambiguation of Verbs by Collocation (DVC) Project	13
2.5. Contrastive Interlanguage Analysis (CIA).....	14
2.5.1. CIA and Interlanguage	15

2.6. Studies Related to CIA, CPA and PDEV	15
2.7. Learner Corpus Research and Error Analysis	17
2.7.1. Error or Innovation?	18

CHAPTER THREE

3. METHODOLOGY	19-26
3.1. Introduction	19
3.2. Data	19
3.2.1. LOCNESS	21
3.2.2. ICLEv3 (International Corpus of Learner English)	21
3.2.2.1. TICLE (Turkish International Corpus of Learner English)	22
3.2.2.2. NICLE (Norwegian International Corpus of Learner English)	22
3.2.2.3. JPICLE (Japanese International Corpus of Learner English)	23
3.3. Data Analysis Tools	23
3.3.1. Sketch Engine	23
3.3.2. The Pattern Dictionary of English Verbs (PDEV)	24
3.4. Research Procedures	24

CHAPTER FOUR

4. FINDINGS AND DISCUSSION	27-47
4.1. Data Analysis	27
4.1.1. Selection of Target Verbs	27
4.1.2. Identifying Patterns	29
4.2. Identifying <i>accept-n</i> Patterns in Target Corpora	32
4.3. Identifying <i>Allow-n</i> Patterns in Target Corpora	34
4.4. Divergency of <i>Accept</i> and <i>Allow</i>	37
4.4.1. Focus on <i>Accept</i> in Target Corpora	38
4.4.2. Focus on <i>Allow</i> in Target Corpora	40
4.5. Discussion of the Research Questions	43
4.5.1. Discussion of the First Research Question	43
4.5.2. Discussion of the Second Research Question	45
4.5.3. Discussion of the Third Research Question	46
4.5.4. Discussion of the Fourth Research Question	47
CONCLUSION	48
REFERENCES	52
APPENDICES	60
CURRICULUM VITAE	68

ÖZET

Bu tez, anadili İngilizce olan ve olmayan öğrencilerin fiil-isim kalıplarını nasıl kullandıklarını araştırmayı amaçlamaktadır. Bu amaç doğrultusunda, bu tez Türk, Norveç ve Japon öğrenenlerin tartışmacı denemelerinde fiil-isim kalıplarının kullanımına odaklanmaktadır. Farklı anadil geçmişlerinden gelen ve anadili İngilizce olan öğrenenlerin tartışmacı denemelerinde fiil-isim kalıplarını nasıl kullandıklarını görmek için derlem-kalıp analizi (CPA) metodolojisi uygulanarak Karşıtsal Aradıl Çözümlemesi (CIA) kullanılmıştır. International Corpus of Learner English'in (ICLE) alt derlemleri olan TICLE, NICLE ve JPICLE öğrenen derlemlerinde hem nicel hem de nitel analizler yapılmış ve bu derlemler anadili İngilizce olan kişilerin makalelerinden oluşan Louvain Corpus of Native English Essays (LOCNESS) ile karşılaştırılmıştır. Hangi fiillerin çalışılacağına karar vermek için PTE Eşdizimlilik listesi referans olarak kullanılmıştır. Fiiller bazı sıkı seçim kriterine göre seçilmiştir. Nicel analize gelince, hedef derlemler arasında istatistiksel olarak anlamlı bir fark olup olmadığını gözlemlemek ve öğrencilerin denemelerinde hedef kalıpları ne ölçüde kullandıklarını görmek için tanıklı dizin satırlarında frekans ve log-olabilirlik analizi yapılmıştır. Tanıklı dizin satırları, fiiller için kalıp açıklamaları sağlayan çevrimiçi bir araç olan İngilizce Fiiller Kalıp Sözlüğü'ndeki (PDEV) kalıplarıyla eşleştirilmiştir. Ayrıca, nitel analiz, öğrenci derlemleri arasında tanıklı dizin satırlarının karşılaştırmalarını içermektedir. Nicel analiz sonucunda, yalnızca *accept* ve *allow* fiilleri ileri analiz için istatistiksel olarak anlamlı bulunmuştur. İki fiil için yapılan analiz, PDEV'de herhangi bir eşleşme bulunmayan bazı olası ıraksak kalıplar göstermiştir. Bu nedenle, ıraksak olup olmadıklarına karar vermek için daha fazla inceleme yapılmıştır. Bu kalıpları kontrol etmek için Grammar Patterns 1: Verbs kitabı ve bu kalıpların anadil verilerinde sıklıklarının olup olmadığını kontrol etmek için İngiliz Ulusal Derlemi (BNC) kullanılmıştır. Bu tezin bulgularına dayanarak, CPA/PDEV'in öğrenen derlemlerine kolaylıkla uygulanabildiği; ancak denemelerinde fiil-isim kalıplarını kullanma konusunda Türk ve Norveçli öğrenenler arasında farklılık, Türk ve Japon öğrenenler arasında ise benzerlik olduğu görülmüştür. Öğrenenlerin eşdizimsel farkındalıklarını artırmak için öğretim materyallerinde CPA yöntemi uygulanarak fiil-isim eşdizimlerinin öğretilmesi önerilebilir.

Anahtar Kelimeler: Öğrenen Derlemi, Fiil-isim Kalıpları, CPA, CIA

ABSTRACT

This thesis aims to investigate how native and non-native learners of English use verb-noun patterns. In line with this purpose, this thesis focuses on the use of verb-noun patterns in the argumentative essays of Turkish, Norwegian and Japanese learners. Contrastive Interlanguage Analysis (CIA) was employed through applying corpus-pattern analysis (CPA) methodology to see how learners from different L1 backgrounds and native speakers of English use verb-noun patterns in their argumentative essays. Both quantitative and qualitative analyses were employed and TICLE, NICLE and JPICLE, sub-corpora of the International Corpus of Learner English (ICLE) were compared to the Louvain Corpus of Native English Essays (LOCNESS) consisting of articles by native speakers of English. Pearson Test of English (PTE) Collocation list was used as a reference for word selection. Verbs were selected based on several strict selection criteria. As for the quantitative analysis, frequency and log-likelihood analysis were done to observe if there was a statistically significant difference across target corpora and to see to what extent learners used the target patterns. The concordance lines were matched with their patterns in the Pattern Dictionary of English Verbs (PDEV) which was an online tool providing pattern descriptions for verbs. Moreover, the qualitative analysis included concordance lines comparisons across three learner corpora. As a result of quantitative analysis, only verbs *accept* and *allow* were found statistically significant for further analysis. The analysis for two verbs showed some possible divergent patterns which did not have any matches in the PDEV. Thus, further examination was done to decide whether they were divergent or not. *Grammar Patterns 1: Verbs* was used to check these patterns and British National Corpus (BNC) was used to check if these patterns had their frequency in native data or not. Based on the findings of this thesis, it was seen that CPA and PDEV were easily applied to learner corpora; however, it was also observed that there was a difference between Turkish and Norwegian learners and similarity between Turkish and Japanese learners for the use of verb-noun patterns in their essays. It could be suggested that learners who have different L1 backgrounds should be taught verb-noun collocations in teaching materials to increase collocational awareness.

Keywords: Learner Corpora, Verb-noun Patterns, CPA, CIA

LIST OF TABLES

Table No	Table Name	Page Nr.
1	Collocation Patterns.....	11
2	Comparison Between the Target Corpora (ICLE and LOCNESS)	20
3	Verbs and Their Raw Frequencies in Four Target Corpora (Verb-Noun Collocations).....	27
4	Frequencies of Verbs in Each Source	29
5	PDEV Entry for Affect in TICLE.....	30
6	Log-Likelihood Values and Over/Under Representations.....	31
7	Statistical Analysis of accept-n Patterns	32
8	PDEV Entries of Accept-N Patterns.....	33
9	Log-likelihood and Over/Under Representations of Patterns For Accept-N	34
10	Statistical Analysis of allow-n Patterns	35
11	PDEV Entries of allow-n Patterns	35
12	Log-likelihood and Over/Under Representations of Patterns for allow-n	36
13	Raw Frequency and LL Values of accept.....	38
14	Pattern Classification of accept in TICLE Instances Not Matched in PDEV	39
15	Pattern Classification of accept in NICLE Instances Not Matched in PDEV	39
16	Pattern Classification of accept in JPICLE Instances not Matched in PDEV	40
17	Raw Frequency and LL Values of allow	41
18	Pattern Classification of allow in TICLE Instances Not Matched in PDEV	42
19	Pattern Classification of allow in NICLE Instances Not Matched in PDEV	42
20	Pattern Classification of allow in JPICLE Instances Not Matched in PDEV	43
21	Verb-Noun Patterns of accept in Target Corpora	44
22	Verb-Noun Patterns of allow in Target Corpora.....	44

LIST OF FIGURES

Figure No	Figure Name	Page Nr.
1	Contrastive Interlanguage Analysis	14
2	Screenshot of the Word Sketch Function in Sketch Engine	24
3	Research Design	26
4	Target <i>Verb-n</i> Distribution	31
5	<i>accept</i> and <i>allow</i> Distribution across Corpora	37
6	Concordance Lines of <i>accept</i> in TICLE which Did not Match Patterns in PDEV	38
7	Concordance Lines of <i>accept</i> in NICLE which Did not Match Patterns in PDEV	39
8	Concordance Lines of <i>accept</i> in JPICLE which Did not Match Patterns in PDEV	40
9	Concordance Lines of <i>allow</i> in TICLE which Did not Match Patterns in PDEV	41
10	Concordance Lines of <i>allow</i> in NICLE which Did not Match Patterns in PDEV	42
11	Concordance Lines of <i>allow</i> in JPICLE which Did not Match Patterns in PDEV	43

LIST OF ABBREVIATIONS

BNC	: British National Corpus
CIA	: Contrastive Interlanguage Analysis
CPA	: Corpus Pattern Analysis
DDL	: Data-Driven Learning
DVC	: Disambiguation of Verbs by Collocation
E1	: English as a First Language
E2	: English as a Second Language
E2F	: English of French Learners
E2G	: English of German Learners
E2J	: English of Japanese Learners
E2S	: English of Swedish Learners
EA	: Error Analysis
ICLE	: International Corpus of Learner English
IL	: Interlanguage
ILV	: Interlanguage Varieties
JPICLE	: Japanese International Corpus of Learner English
L2	: Second Language
LAD	: Language Acquisition Device
LOCNESS	: Louvain Corpus of Native English Essays
NICLE	: Norwegian International Corpus of Learner English
NL	: Native Language
NNS	: Non-Native Speaker
NS	: Native Speaker
PDEV	: Pattern Dictionary of English Verbs
PG	: Pattern Grammar
PTE	: Pearson Test of English
RLV	: Reference Language Varieties
SLA	: Second Language Acquisition
TICLE	: Turkish International Corpus of Learner English
TNE	: Theory of Norms and Exploitations

INTRODUCTION

“The study of language is the study of patterns.”

J. Sinclair

English functions as a lingua franca in diverse contexts from business to academia where the linguistic features in learner language offer insights into the linguistic differences (Flowerdew, 2015: 466; Jeckins et al., 2011: 282; Seidlhofer, 2007: 137). Learning English as a foreign language has also become one of the goals of students and understanding the features of learner language is one of the fundamental elements of language learning process. In this context, collocations play a crucial role for language learners to produce more natural and fluent language. Corpus-based studies have potential to provide an effective method for teaching learner collocations based on authentic language data. However, the structural diversity in native and target language may make learners experience difficulties in comprehension of collocations.

Over the past two decades, corpus-based studies have increased to give learners a sense of awareness of authentic language due to its real language data applied to the language learning process (McEnery et al., 2006: 4; McEnery & Brezina, 2022: 6). As a part of a language, multi-word combinations that are used together in a language are closely related in meaning and they form a unit of meaning. The correct use of collocations is one of the basics for a more natural language production (Ellis et al., 2015; Erman et al., 2016: 111; Gablasova et al., 2017: 159); on the other hand, the incorrect use of collocations might lead to grammatical errors or semantic shifts (Allen, 2011; Laufer & Waldman, 2011: 648; Nesselhauf, 2005). It is essential to understand and apply the structural features of the language correctly; however, the native language of L2 learners might affect their use of collocations. Up to now, learner corpus studies involved comparing findings of learner corpora analysis with the conclusions derived from the native corpus analysis to identify collocational patterns and errors in the over-underuse of these collocations in learner data.

On the other hand, learner corpus studies have so far been contrastively conducted to identify the interlanguage differences between learners. This area of learner corpus research is Contrastive Interlanguage Analysis (CIA), which determines the differences between learners' native and target languages (Granger, 2015: 8; Granger & Tyson, 1996: 18). The interlanguage analysis helps to examine whether learners produce incorrect collocations due to their L1 or not and help learners understand the grammatical differences to improve their correct use of collocations and reduce errors. Learning collocations contribute to learners' language skills; therefore, the International

Corpus of Learner English (ICLE) provides a valuable database for contrastive studies. ICLE can make a great source of L2 learner analysis as its latest version consists of 25 different language backgrounds (Granger et al., 2020: 7).

Furthermore, Corpus Pattern Analysis (CPA) is recently used in corpus studies to investigate the grammatical structures and patterns of language using corpus evidence (Hanks, 2004: 87). CPA performs statistical and linguistic analysis to identify the co-occurring structures in sample data. Especially in verb-noun collocation studies, CPA is used to detect which verbs are frequently used with nouns and which grammatical patterns these usages conform to. Moreover, the Pattern Dictionary of English Verbs (PDEV) is a source that shows with which patterns various verbs are used together. This dictionary presents specific patterns of verbs and their implications based on native data (British National Corpus data) to help language learners to comprehend and practice correct verb patterns.

All in all, the aim of corpus linguistics is to provide descriptions of authentic language to be used by language learners to explore how language is used in context. Regarding this, the thesis aims to contrastively examine the verb-noun collocation pattern use of Turkish, Norwegian and Japanese learners along with the native speakers of English through the use of the PDEV.

CHAPTER ONE

1. FRAMEWORK OF THE STUDY

1.1. Introduction

The Corpus Pattern Analysis (henceforth CPA) is developed by Patrick Hanks as a corpus-based technique based on the idea that “meanings can be associated with phraseological patterns as well as or instead of words in isolation” (Hanks, 2012: 68). CPA underlies the idea that language consists of patterns which are more related to the language rather than the meaning of a word (El Maarouf, 2013: 121). The syntactic and syntagmatic relation between the verbs can also be explained through this method (Cinkova & Hanks, 2010: 1). It is important during CPA applications that analyses of concordance lines of specific words come before reaching the meaning of these words (Rees, 2021: 35). As a practical outcome of the CPA method, PDEV serves as a worthwhile tool for describing the utilisation of verb patterns in both native and learner corpora. The main purpose of the PDEV is to reveal the collocational patterns of English verbs looking at the concordance lines gathered from the 50-million-word subcorpus of British National Corpus (BNC) (Rees, 2021: 36). The theoretical value of the PDEV is derived from corpus-based CPA patterns of English verbs with their implications offering semantic annotation of verbs and their collocations (Rees, 2018: 24). It emerged as a result of systematic analyses of English verbs of BNC from norms which refers to the normal patterns of verbs to exploitations referring to the creative use of words; then, meanings and patterns are related (Hanks, 2013: 404-405).

In the research background, many learner corpora studies use Contrastive Interlanguage Analysis (CIA) methodology which comprises both native and non-native data to compare. Granger (1996: 43) proposed CIA approach as a combination of contrastive analysis and interlanguage analysis to investigate the similarities and differences of non-native production of English with native speakers of English. The linguistic variations in learner corpora studies also led this study to apply CIA method. However, it was seen that there were limited number of research which were using the PDEV. Rees (2018) studied the vocabulary selections, differences and use across disciplines to apply the CPA method. The PDEV, in this study, is used as a framework as it provides the semantic annotation of verbs accurately. On the other hand, Alqarni (2018) specifically studied the pedagogical potential of the PDEV by contrasting learner corpora with a native corpus. The study proposed that utilising the PDEV may enhance learner awareness by

revealing the divergent patterns in target corpora. It should also be stated that this current thesis is inspired by Alqarni's (2018) study.

The objective of this thesis is to investigate the verb-noun patterns in learner corpora from different language backgrounds. Three learner corpora (TICLE, JPICLE and NICLE) were selected to extract the target patterns and a native corpus (LOCNESS) was used to compare the way how native and non-native learners used verb-noun patterns.

1.2. Purpose of the Study

The purpose of this thesis is to scrutinize the use of verb-noun patterns in three learner corpora with diverse backgrounds and compare them to a native corpus. It was anticipated that the analyses demonstrated the use of verb-noun patterns of non-native learners and their possible correspondence with native learners. For this investigation, two corpora were used. The first is the International Corpus of Learner English (ICLE) and the other is the Louvain Corpus of Native English Essays (LOCNESS). More specifically, three sub-corpora of ICLE are used to analyse. The initial one is TICLE which functions as the Turkish sub-component of ICLE. The second is NICLE which represents the Norwegian sub-component of ICLE. Lastly, JPICLE functions as a Japanese sub-component of ICLE. The extracted concordance lines of the verb-noun collocations are used to analyse their patterns with the PDEV.

Learner corpora are the collections of L2 learners' language data that are designed based on strict criteria including learner variables such as gender, age, education and their linguistic backgrounds (Granger, 2019: 3). To identify the differences between these variables, CIA is one of the most effective ways as it provides comparisons of learner variables with a reference corpus (Granger, 2015: 8; 2019: 3). In this study, ICLE is chosen as a learner corpus which consists of learner data from different backgrounds such as Turkish, Norwegian, and Japanese learner data used in this study. The purpose of the selection of these language backgrounds is that mainly the researcher of this thesis is the Turkish learner of English which aims to show the use of verb patterns by Turkish learners with a contrast to other language backgrounds. JPICLE is chosen as it is also from the same language family as Turkish language and they both share the same language structure (subject- object- verb). NICLE is chosen as it shows more proficient language user profile than TICLE and JPICLE and English in Norway is part of daily life which can be their second language (Granger et al, 2020: 168).

1.3. Significance of the Study

CPA, as the name suggests, derives corpus-based patterns which are represented with frequent occurrence in relation to their general meaning (Rees, 2018: 26). Hanks (2004: 88) stated

that “in CPA, the ‘meaning’ of a pattern is expressed as a set of basic implicatures”. Successful implementation of CPA led to the emergence of the PDEV as a new technique to describe the thousands of verbs in English with their patterns and their implications.

The purpose of this thesis is to provide a deeper understanding of the use of verb-noun patterns of Turkish, Norwegian and Japanese learners of English in comparison to native English speakers. It intends to give a comprehensive overview about whether these learners exhibit similar use of verb-noun patterns among themselves and native speakers. To investigate the overuse and underuse of the target verbs in three corpora, the CPA method, and its productive outcome the Pattern Dictionary of English Verbs (PDEV) was used. The significance of this thesis stems from the machine-readable nature of the PDEV (Hanks & Pustejovsky, 2005: 64), which provides an easily accessible way to analyse the verb pattern in learner corpora. In an attempt to examine the learner corpora, this thesis used a newly developed source that has not been widely investigated. It is hoped to make an extensive analysis of the use of verb-noun patterns across learner corpora from different backgrounds with a comparison to a native corpus. Another important aspect of this thesis is that it investigates whether only Turkish learners of English use verb-noun patterns differently/similarly when compared to Norwegian and Japanese learners of English. For this investigation, a two-way analysis using LOCNESS and PDEV helped us to assess verb-noun usage patterns in three corpora with different L1 backgrounds. It is also essential to study verb patterns as it provides semantic information showing “the grammar patterns are less informative than the PDEV entries” (Hunston & Su, 2017: 2). The thesis findings are also expected to show the strength and weaknesses of learner language while using verb-noun patterns, which give insights for future collocation teaching in language classrooms. On the other hand, the thesis deals with the erroneous use of these patterns, which may affect the future of English learning. It is expected for this thesis to fill the gap in the literature as being the pioneer by applying the PDEV to contrastive examination.

1.4. Statement of the Problem

Firth (1957) introduced the term ‘collocation’ to describe the companion of words. Sinclair (1991: 109) supported the idea by defining collocations as the frequent occurrence of multiword expressions. Collocations are word combinations having a grammatical pattern which shares meaningful associations with each other (Nesselhauf, 2005). As collocations are defined in literature as “the most natural way of saying something in English” (McCarthy & O’Dell, 2005: 6), it is important to teach students verb-noun collocations through awareness-raising activities (Gençer, 2004), DDL applications, explicit instruction or combined method (both DDL and explicit instruction) (Akıncı, 2009). In this respect, collocations are an inseparable part of producing natural language patterns and being proficient in English. It seems that there is still a need for investigating

learners' verb-noun collocations with PDEV, a different and new method within the framework CPA.

The aim of this thesis was to examine the similarity and differences in the use of verb-noun patterns by Turkish EFL learners when they are compared to native speakers of English and other English learners who are sharing different language backgrounds. It was important to observe different language backgrounds by using interlanguage comparisons to detect the effects of individual differences and experiences of language learners on the use of verb-noun patterns. It was aimed to investigate how the use of verb-noun patterns varies across different groups, using LOCNESS and PDEV as secondary sources which are generally used as comprehensive sources of data for linguistic analysis.

1.5. Research Questions

To address the aforementioned objectives, this thesis was designed to answer the following research questions below:

1. What are the similarities and differences between verb patterns used by learners in TICLE, NICLE, JPICLE and the LOCNESS corpora?
2. What are the probable divergent patterns in the use of verb-noun patterns in the learner corpora?
3. How can CPA be used to describe verb-noun patterns in learner corpora with different L1 background?
4. To what extent is the use of PDEV for the description and identification of target divergent patterns possible in learner corpus analysis?

1.6. Organisation of the Thesis

This thesis aims to investigate verb-noun patterns used by learner corpora from different backgrounds. Thus, this thesis is carried out by using the Pattern Dictionary of English Verbs (PDEV) to determine the patterns in target corpora. This study consists of five chapters.

Chapter 1, Framework of the Study outlines the primary reasons why this thesis was studied and also underscores the importance of the thesis. Furthermore, it explicates the importance and the purpose of the study including the research questions.

Chapter 2, Literature Review is the review of the related literature about collocation, pattern grammar, corpus pattern analysis, semantic preference, verb pattern and learner corpus research. This chapter also helps to identify the gaps in related literature, to analyse the results of the literature and to provide potential directions for future research.

Chapter 3, Methodology outlines the employed methodology. It involves data, data analysis tools and research steps. This chapter provides a detailed description about the research procedures.

Chapter 4, Findings and Discussion reveals the data analyses. It also summarises and discusses the findings one by one within the tables and figures. This chapter also includes how the findings can contribute to practical applications, limitations and suggestions for future studies. The implications of the findings are analysed in depth.

Chapter 5, Conclusion summarises the main findings and also provides implications for future studies while commenting on the findings based on the answers given to the research questions. This chapter is a concluding section of the thesis and emphasizes the significance and contribution of the thesis.



CHAPTER TWO

2. LITERATURE REVIEW

2.1. Introduction

This chapter presents the contextual background to provide additional knowledge to the central concept of this thesis which is corpus pattern analysis (CPA). This chapter primarily provides definitions and approaches in order to make the concepts comprehensible. In the following sections, the notion of pattern - a key concept of pattern grammar- which is also a framework emphasising the importance of patterns in language acquisition and use was discussed. Later, Corpus Pattern Analysis was given in detail a valuable method for investigating the language learning process and the language organisation in a different context. Further, it demonstrated how PDEV is used in corpus-based English verb studies. Finally, the learner corpus research section was given along with the literature based on learner errors.

2.2. The Notion of Pattern

‘Pattern Grammar’ (PG) is an approach to English grammar that prioritises the patterns of specific words gathered from the concordance lines (Francis, 1993; Hunston & Francis, 2000; Hunston, 2015; Sinclair, 1991, 2004). Unlike traditional grammar which only provides knowledge about the language, Hornby (1975) emphasized the importance of patterns as they signify how language can be used to gather words in proper order. The roots of the notion of ‘pattern’ dates back to Hornby’s (1954) work and Sinclair (1991: 113) theorised the term based on the correlation between words and their structures occurring together. The notion was further developed as pattern grammar (Hunston & Francis, 2000).

Hornby’s (1954) work *A Guide to Patterns and Usage in English* is the precursor to pattern studies. His descriptions for patterns are intended to explain 25 verb patterns, 4 noun patterns, 3 adjective patterns and adverbs. It is also pointed out in this book that verb patterns are the most significant patterns that “when the learner learns the meaning of the adjective *anxious*, he should also learn its patterns: ‘anxious about his son’s health’, ‘anxious for news’, ‘anxious (=eager) to start’.” (Hornby, 1954: 6). However, concerning the descriptions of patterns instead of analysing these patterns prevents him to describe structure (Hunston & Francis, 2000: 4). Although it was a

good attempt to observe English patterns, technological advancements enhanced the quality of works through corpora and fulfilled the missing points of his works (Hunston & Francis, 2000: 7).

Sinclair in his book *Corpus Concordance Collocation* (1991: 4) stated that “the ability to examine large text corpora in a systematic manner allows access to quality of evidence that has not been available before”. The connection between words and grammar entails the emergence of a new theory which is proposed by Sinclair (1991). It is essential to account for two principles: the open-choice principle and the idiom principle.

According to the open-choice principle, there are many *slots* to be *filled* with the choice of lexical items which is the decision of the language user to grammatically complete these slots (Sinclair, 1991: 109; Barnbrook, 2007: 188). The contrast between the idiom principle and the open-choice principle serves to the interpretation of language. Therefore, the open-choice principle put grammatical pressure on language and is not enough to explain language because words do not occur randomly in texts. Sinclair (1991: 109-110) states that every language is based on the open-choice principle; however, the idiom principle -at this point- emerged to fill the gap in which the open-choice principle falls behind in explaining consecutive word choices in a text.

The Idiom principle includes collocations which consist of words which are frequently occurred together and change meaning according to their combinations. It is also defined as “a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments” (Sinclair, 1991: 110). Furthermore, the idiom principle separates *terminological tendency* from the *phraseological tendency* that the former refers to the fixed meanings of words belonging to the world and the latter pertains to the word choices based on their collocations making meanings together (Sinclair, 2004: 29; Hanks, 2013: 399).

The relationship between meaning and pattern was further developed by Francis (1993, 1995) and Hunston and Francis (1998, 2000). It was seen that “certain patterns ‘select’ words of particular meanings” (Hunston & Francis, 2000: 29). Based on this, the concept of ‘Pattern Grammar’ (PG) was developed as a corpus-based approach that explains the importance of patterns and enables English learners to use the language properly (Francis et al., 1996; Hunston, 2002). The outcomes of this concept appeared in *Collins COBUILD English Dictionary* (Francis et al., 1996). The patterns of more than 4,000 words are examined in this dictionary. Patterns of words are identified based on their frequency, dependency on word choice and creating meaning (Hunston & Francis, 2000: 8).

Pattern Grammar is a useful approach for practical linguistic investigations; however, PG cannot be used to make a distinction on collocations which can count as a pattern (Hanks, 2013: 6).

According to pattern grammar, for example, “there is no distinction between *firing a gun* and firing *an employee* since they both transitive uses of the verb *fire*, with pattern V_n” (Hanks, 2013: 6). As a result of such limitation, Corpus Pattern Analysis (CPA) is introduced as a new method.

2.3. Corpus Pattern Analysis (CPA)

Corpus Pattern Analysis (CPA) was developed by Hanks (2004: 87) to analyse English word patterns systematically by “mapping meaning onto words in text”. CPA fundamentally focuses on identifying patterns of actual use of words and associating implicatures with patterns of these words (Cinkova & Hanks, 2010). The CPA technique is underpinned by the *Theory of Norms and Exploitations (TNE)* (Hanks, 2013: 404). It is a ‘*double helix*’ approach and Hanks (2013: 411) defined this approach as “the set of rules that govern normal, conventional use of words is intertwined with a second-order set of rules that govern how those norms can be exploited and that contribute very largely to the phenomenon of language change”. It is possible to better understand the line between norm and exploitations by using large corpora and statistical evidence (Hanks & Može, 2019). Any pattern may be counted as both norm or exploitation; however, the normality of these patterns is distinguished with lexical, syntactic or semantic-type alternations which make them more or less normal (Hanks, 2013: 411). Moreover, what makes a norm exploitation is the recurrent nature of exploitation which rejects the rules of a normal pattern (Hanks, 2013: Franklin, 2020).

In CPA, patterns are identified by grouping concordance lines according to their lexical, semantic, and syntactic features. Words may not have meanings alone, but they contribute to the meaning in a context (Hanks, 2004: 88; Hanks & Pustejovsky, 2005: 64). CPA is based on the investigation of verb patterns in English. Also, the CPA of verbs is important because verbs are the base of a sentence, and their patterns are different from nouns in terms of the fact that noun patterns can be semi-automatically investigated through the accurate analysis of verb patterns (Hanks, 2004: 92). As an outcome of CPA, *Pattern Dictionary of English Verbs* was developed as a new kind of online pattern dictionary to define the verb patterns and their meanings that emerge from the patterns of each verb (Hanks, 2013; Hanks & Mu, 2020).

On the other hand, CPA is underpinned by Sinclair’s ‘units of meaning’ concept which includes collocation, colligation, and semantic preference. The concept was developed by Sinclair (1991, 2004: 3) and he defined it as “choices of lexical patterns entail patterns of meaning and even more precisely every distinct sense of a word is associated with a distinction in form.” (p.3).

2.3.1. Collocation and CPA

Collocations have ubiquitous nature in both written and spoken English and they emerged as a result of natural-sounding word combinations in the language (Oxford Collocations Dictionary for Students of English, 2002). They are the word groups that are expected to naturally come together in the same context (Firth, 1957). On the other hand, Sinclair (1991: 170) defines collocations as “the occurrence of two or more words within a short space of each other in a text”. Collocations can be seen in several ways and McCarthy and O’Dell (2005; 2017) classified collocational patterns into six types (Table1).

Table 1: Collocation Patterns

Type	Example
Adjectives and nouns	major problem
Nouns and verbs	companies merged
Noun + noun	a sense of pride
Verbs and prepositional expressions	burst into tears
Verbs and adverbs	smiled proudly
Adverbs and adjectives	happily married

Source: McCarthy & O’Dell (2005; 2017)

There are also many verb-noun collocations which are associated with various patterns in CPA. It provides us with corpus evidence that enables us to describe the patterns of the normal use of verbs and their collocations (Hanks & Može, 2019). In PDEV, some verb patterns include collocates which means they always accompany together. For example, the PDEV entry for the verb *open* contains fifty-six patterns. Of these patterns, there are some patterns including different collocates: *door* and *the way*. Each of these nouns has different meanings when collocating with *open*. Some examples are below:

(2.1) Tamsin **opens the door** and smiles. (BNC)

(2.1.1) [[Human]] open {door | window | gate}

(2.1.2) [[Human]] moves a {door | window | ...} so that there is a gap between the {door | window | ...} and its frame, allowing access, vision, or the flow of air

The meanings of the verb *open* in (2.1) and (2.2) are different. (2.1.1) and (2.2.1) show the PDEV entries for patterns of the concordance lines which are gathered from BNC. Also, (2.1.2) and (2.2.2) are the implicatures of the patterns which are roughly equal to the dictionary meanings of the entries. In (2.1), a person moves the door to allow access, on the other hand, the eventuality provides innovation to remove an obstacle preventing a future occasion.

(2.2) The committee's recommendations **open the way** for a secure future. (BNC)

(2.2.1) [[Human | Institution | Eventuality 1]] open {the way} [for | to [[Eventuality 2 = Desirable]]]

(2.2.2) [[Human | Institution | Eventuality 1]] removes the obstacles that have been preventing [[Eventuality 2 = Desirable]] from happening

2.3.2. Colligation and CPA

According to the definition of Hoey (2005), colligation refers to the grammatical accompanying of words that occurred or were avoided in a certain form in a sentence. Hanks (2004: 93) indicated the importance of colligation as the sub-valency features may even change the meaning itself with or without the absence of grammatical information. This means that valency and sub-valency feature share the same importance in the meaning of verb patterns. Hanks (2004: 87) provides his claim with an example: there is a difference in meaning between *take place* and *take his place*. The latter has a sub-valency feature '*his*' that distinguishes the meaning of a pattern from another.

2.3.3. Semantic Preference and CPA

Hanks (2013) suggested the term semantic types as a similar term to semantic preference which refers to the semantically co-occurred words. In CPA, an ontology of semantic types is employed to investigate the nature of general language rather than genre-specific language (Rees, 2018: 211). The PDEV ontology consists of 225 semantic types which "form the basis of the semantic strings from patterns and link them to verbs" (Bradbury & El Maarouf, 2013). A lexical set is constituted with a combination of words and this lexical set is unified by semantic types which are sequentially aligned as ontology (Hanks, 2013). For example, *policemen*, *teachers*, and *journalists* can be generalised by the semantic type [[Human]] while *the university*, *the office*, and *the hospital* can be generalised by the semantic type [[Institution]].

2.4. Pattern Dictionary of English Verbs (PDEV)

Hanks (2013), in his TNE theory, asserted that the language-related problems can be solved if the words and their meanings are associated each other with their patterns. His TNE theory became the theoretical framework for CPA which aims to how English verbs can be identified. As a result, the Pattern Dictionary of English Verbs (PDEV) emerged as an online pattern dictionary consisting of verbs with their phraseological patterns; implicatures of each pattern; random samples for each pattern taken from the British National Corpus (BNC); and ontology of semantic types (Hanks, 2013; El Maarouf et al., 2013; Popescu et al., 2014). It is an ongoing project of CPA which currently has 1,691 completed verb entries. Verb patterns in PDEV involve both valency and

collocations and the latter are grouped based on their semantic types (Hanks, 2012; Popescu et al., 2014).

In PDEV, the BNC50 corpus was used as a reference from which sample concordance lines were chosen. The BNC50 sub-corpus consists of 50 million words excluding the spoken part of BNC as spoken data do not “conform to a theoretical prediction” (Hanks, 2004: 88) and it is difficult to analyse and interpret the spoken data.

2.4.1. Identifying Verb Patterns

Patterns are investigated by collecting and grouping some concordance lines according to their lexical, semantic and syntactic features. In CPA, all patterns “consist of a valency structure ... together with sets of preferred collocations” (Hanks, 2013: 92). Patterns can be either norms or exploitations depending on their normal or rare use respectively. In PDEV, verb patterns are established according to their semantic types which form the PDEV *ontology*. The ontology is divided into five main categories: Group, Eventuality, Entity, Part, and Property. There are sub-categories of these main categories. For example, the Part category has seven sub-categories: Document Part, Language Part, Movie Part, Physical Object Part, Recording Part, and Speech Act Part. These sub-categories have also sub-categories. Also, semantic types are given in double brackets. To exemplify, the verb *allow* has pattern 1 which is listed in PDEV as

(2.4) [[Human 1]] or [[Institution 1]] or [[Eventuality]] **allow** [[Human 2]] or [[Institution 2]] or [[Self]] [to+INF]

(2.5) [[Human 1]] or [[Institution 1]] or [[Eventuality]] gives **allow** [[Human 2]] or [[Institution 2]] or [[Self]] permission or the opportunity to have [[Privilege]]

In the given example, (2.4) shows the pattern of the verb *allow* and (2.5) is the implicature of the pattern which roughly represents the dictionary meaning. Implicatures indicate the pragmatic meanings of verbs and pattern descriptions (Hanks, 2013).

2.4.2. The Disambiguation of Verbs by Collocation (DVC) Project

The Disambiguation of Verbs by Collocation (DVC) project is an outcome of CPA aiming to improve PDEV on a larger scale by examining the link between the verb patterns and their meanings (El Maarouf, et al., 2014a; El Maarouf et al, 2014b). The main purpose of the DVC project is to investigate the normal and collocational use of verb patterns; thus, 250 concordance lines for each verb were examined and linked to the patterns to which they are related (Hanks & Bradbury, 2013). It is also a outcome of the relationship between lexicographers and computer

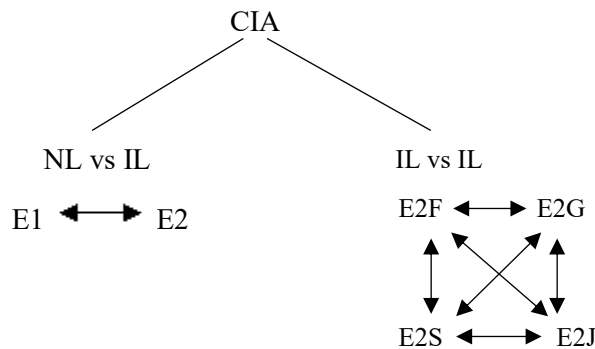
scientists that need lexicological knowledge rather than scientific theory to associate verbs and their meanings using ontology (Hanks & Bradbury, 2013; El Maarouf et al, 2014).

2.5. Contrastive Interlanguage Analysis (CIA)

Linguists have engaged in comparing different language backgrounds in their studies to elicit the potential differences and similarities in second language learning. A new framework has been developed alongside the compilation of ICLE for learner corpus research and it is Contrastive Interlanguage Analysis (CIA) (Granger, 1996: 43). It emerged as “the core of ICLE project whose main aim is to uncover of ‘foreign-soundingness’ in learner writing” (Granger, 1996: 43). Contrary to the other contrastive analysis methods, CIA both qualitatively and quantitatively compares the diversity of the same language to evaluate the differences between L2 English from L1 English (Hasselgård & Johansson, 2011). It is used to compare native and learner corpus to reveal the low and high frequent lexical items and structures in target corpora; however, it should be noted that the texts whose comparisons are made should belong to the same genre not to cause unreliability (Yıldız & Turan, 2021).

CIA comparisons can be interpreted in two sections. Initially, the comparison of learner language with native language (NL) detects the ‘overuse’ and ‘underuse’ of linguistic features. On the other hand, the second type of this comparison consists of different interlanguages (IL) of the same language. In Figure 1, the representation of two types of CIA is given.

Figure 1: Contrastive Interlanguage Analysis



Source: Granger, 1996, 44

According to Granger’s (1996) classification of the CIA, there are two types of CIA. Initially, a comparison of native language (NL) and interlanguage (IL) is made in the International Corpus of Learner English (ICLE) by using English as a first language (E1) and as a foreign language (E2). As NL/IL comparisons need a reference corpus, the Louvain Corpus of Native English Essays (LOCNESS) may serve as a source of native data to compare interlanguage corpus such Turkish-

ICLE (TICLE). Conversely, the second comparison category pertains to the diverse interlanguages of the same or different languages. The comparison may involve English of French learners (E2F), German learners (E2G), Japanese learners (E2J), and Swedish learners (E2S). The main purpose of the interlanguage comparison is to better understand what differs between non-native backgrounds (Granger, 1996: 45; 1998; 2015). Furthermore, Granger (2015: 17) reappraised CIA promoting the ‘varieties’ term as reference language varieties (RLV) instead of native language (NL) and interlanguage varieties (ILV) instead of interlanguage (IL). In the first part of the division, diatypic and dialectal variables are important for RLV comparisons. Secondly, task and learner variables provide us to learn further information about learner data.

2.5.1. CIA and Interlanguage

CIA is underpinned by the theory of *interlanguage* which is defined as a naturally occurring language based on the second language learners’ perceptions which share the characteristics of their native and second language (Selinker, 1972; Lakshmanan & Selinker, 2001; Tarone, 2006). Second language learners and native speakers share different linguistic characteristics that may indicate the learners’ inability to produce the target language fluently and proficiently. Learners experience challenging periods in their learning process by making mistakes which may be later fossilised. Especially adult learners of the target language are considered unsuccessful when they are compared to children because children can employ the language acquisition device (LAD); thus they generally tend to avoid errors and fossilisations (Tarone, 2006). Based on this idea, Selinker (1992) proposed the latent psychological structure of adult learners which consists of five cognitive principles which consist of a) language transfer, b) transfer of training, c) strategies of second language learning, d) strategies of second language communication, e) overgeneralisation of target language rules. These principles respectively refer to transferring knowledge from one language to another; employing the acquired competence while the language learning process; techniques used for acquiring a language and language communication; and applying general language rules to produce in L2 English. All in all, corpus-based contrastive analyses offer linguists the chance to investigate the language development of learners regarding their linguistic differences in interlanguage.

2.6. Studies Related to CIA, CPA and PDEV

Native and learner language constitute the learner corpus research. Learner corpora are compared to a native corpus to investigate the similarities and differences between the linguistic features of learners from different language backgrounds (Gilquin et al, 2007). The study of learner corpora underlies the Contrastive Interlanguage Analysis (CIA) which consists of native speaker (NS) and non-native speaker (NNS) or learner data (L2). Altenberg and Granger (2001) investigated the use of high frequency verbs focusing on the verb *make* by Swedish and French

learners with a contrast to native data (LOCNESS). They found that learners both underuse the high frequency verbs such as *make* and produce incorrect usages of these verbs. Another finding of their research was that Swedish learners overused *make* + adjective structures rather than French learners because English and Swedish shared a similarity in this structure.

Nasselhauf (2003) examined the use of verb-noun collocations by advanced German learners of English based on the data German sub-corpus of ICLE in comparison with *Oxford Advanced Learner's Dictionary*, *Collins COBUILD English Dictionary*, *The BBI Dictionary of English Word Combinations*, *the Oxford Dictionary of Current Idiomatic English*, and texts from British National Corpus (BNC). Findings showed that German EFL learners had difficulty in the production of collocations due to the influence of their L1.

Gilquin (2012) investigated ICLE learners' use of causative constructions, especially [X MAKE Y V_{inf}] and [X MAKE Y V_{pp}] in comparison with BNCw data. The findings of the study showed that learners showed tendency to use unidiomatic causative constructions which were less occurred in native data (BNCw).

Babanoğlu (2012) also used CIA method to investigate Turkish learners' use of adverbial connectors in their L2 writings. Turkish, Spanish, and Japanese sub-corpora of ICLE and a native corpus (LOCNESS) were used for contrastive analyses. The findings indicated that Turkish learners of English overused these connectors; especially, resultive connectors such as *so* is overused by Turkish learners while native speakers were mostly using contrastive connectors such as *however*.

There are also some studies that are using CPA method and the PDEV compared to other corpus studies. Initially, Rees (2018) used CPA to investigate the differences in meaning and the use of Academic Word List (AWL) between various academic disciplines (History, Management, and Microbiology). The findings of the study showed that the use of specific verbs in specific discipline was different from other disciplines which means AWL has discipline specific meanings based on their context.

Alqarni (2019) used the PDEV to investigate its potential to describe learner corpora through the use of patterns and their meanings. To apply CPA method, Swedish and Chinese sub-corpus of ICLE as learner corpora and LOCNESS as native corpus were used. In this study, patterns of 16 verbs were analysed and the verb *allow* was found the most noticeable for further investigation. The findings of the study also showed that the PDEV can be successfully used for describing learners' pattern choices.

Franklin (2020) applied CPA method to analyse the use of “killing” terms. “People, Products, Pests and Pets” (PPPP) corpus which consisted of animal-related texts was used to examine the animal discourse by using PDEV to decide how killing verbs were used. At the end, CPA was found a practical way for discourse analysis.

Overall, this thesis aims to apply Hanks’ CPA method to describe learner data (TICLE, NICLE and JPICLE) from different backgrounds using the Pattern Dictionary of English Verbs as an outcome of this method.

2.7. Learner Corpus Research and Error Analysis

Making errors and mistakes is a *sine qua non* in learner language in their language acquisition process (Corder, 1982; Kirkgöz, 2010; Can, 2017). However, the difference between error and mistake is that the former refers to the interlanguage ability of learners with a lack of linguistic knowledge while the latter occurs accidentally due to lack of attention which incorrectly affects language (Richard & Schmidt, 2002; Brown, 2007). Learning a new language without making any errors is far from the reality in the sense that there may be some distinctive linguistic differences between L1 and L2. Thus, it is inevitable for learner corpus research to focus on learner errors made in both language comprehension and production (Ellis, 1994). These errors are examined with *error analysis* (EA) which is a technique to investigate the possible reasons causing learner errors and to classify the error types that occurred during the second or foreign language learning process (Richard & Schmidt, 2002). It is important to make an error analysis as it “provides to the researcher evidence of how language is learned or acquired, what strategies or procedures the learner is employing in the discovery of the language” (Corder, 1967: 167).

The source of learner errors may stem from *interlingual* and *intralingual errors* (Brown, 2007) and has both *diagnostic* and *prognostic* purposes (Corder, 1967). *Interlingual errors* emerge because of first language interference which refers to the unfamiliarity of the learners towards the second language until they familiar with the target language system; on the other hand, *intralingual errors* refers to the misinterpretation of the rules of the target language itself such as overgeneralisation (Brown, 2007; Can, 2018; Kirkgöz, 2010).

On the one hand, three aspects of learner errors were summarised by Corder (1967) as 1) its pedagogical benefits to language teachers showing what learners have and have not acquired so far; 2) its evidence for language learning process and 3) its opportunity for learners to explore the target language’s rules. On the other hand, Ellis (1994) mentioned four steps for EA: 1) collecting data of learner language, 2) identification of the errors in collected data, 3) description of the errors and 4) psycholinguistic explanation for the errors. EA, at this point, is a useful method for the analysis of learner language which is the convergency of corpus linguistics and EA. EA makes sense together

with corpus linguistics which provides the data to be analysed and “both of them holding their own rules but making mutual contribution in order to offer results that will enrich SLA theories and language teaching” (López, 2009: 677). However, it mainly focuses on the learners’ error which limits the learner language with the wrong language use rather than the correct use of language (Ellis & Barkhuizen, 2005). Contrastive Interlanguage Analysis (CIA) and Error Analysis (EA) has a significant role in explaining the learner language production in Second Language Acquisition (SLA). Both approaches cannot be dissociated from corpus linguistics, consisting of learner data. ICLE, in this point, provides a large-scale data to analyse learner English.

2.7.1. Error or Innovation?

Generally, non-native features in interlanguage atmosphere are interpreted as error. However, it is defined as innovation if native speakers of language produce such features. In that case, we can assumed that non-native speakers of a language can also make innovations like natives (Edwards & Laporte, 2015). Corpus-based research enable us to investigate not only the erroneous use of linguistics features but also over-underuse of these features by analysing the production of learners (Nesselhauf, 2003; Salazar, 2014; Schneider & Gilquin, 2016). Edwards and Lange (2016) connected innovation with *structural nativisation* which is defined by Schnedier (2007) as the development of distinct linguistic patterns that are unique to certain ESL countries and observed in certain locations. Collocation ratio technique was developed by Schneider and Gilquin (2016) aiming to identify innovations such as the use of verb + preposition and adjective + preposition between EFL and ESL learners. They used ICLE as EFL data and the International Corpus of English (ICE) as ESL data which shows there is no conventional difference between EFL and ESL context. Alqarni (2018) used the notion of *divergence* instead of using errors that referred to the distinct patterns with specific verbs. In this thesis, the patterns which occurs with the incorrect verbs which are used with correct patterns of right verb was called as *divergent*. This study also aims to analyse the divergent usages of verb patterns by learners in TICLE, NICLE, JPICLE sub-corpora.

CHAPTER THREE

3. METHODOLOGY

3.1. Introduction

This chapter describes the data and the methodological approach and initially introduces the corpora used in this study. As the study is concerned with Corpus Pattern Analysis (CPA), the steps taken in the analysis and justification for the selection of the target corpora are given in detail. Moreover, the research tool and procedure for verb selection are presented. The comparability table of four target corpora is also given to show whether the corpora are suitable to compare.

3.2. Data

Considering the aims of this thesis, the selection of the corpus was important. As the main focus was on the linguistic data of non-native English learners, using learner corpus “that is immediately related to the language classroom...a collection of the writing or speech of learners” was important (McEnery et al., 2006: 65). Learner corpora, which consist of linguistic data collected from non-native language learners, provide a wide range of data to identify L2-based features (Gilquin et al., 2007). ICLE was both the most popular and suitable corpora that could be used in this study. It is important to follow a strict design criterion for the compilation of a corpus because the heterogeneous nature of learner language may cause misinterpretation of the data (Granger et al., 2020: 3). The compilation of the International Corpus of Learner English (ICLE) is based on some strict design criteria regarding “the importance of collecting well-defined samples of learner language so that clear statement can be made” (Ellis, 1994: 49). Variables for setting the ICLE project are twofold: learner variables cover the young university learners of English with higher proficiency levels and language variables cover the text selection which consists of argumentative essays written in academic format (Granger et al., 2020: 3-4). ICLE is chosen as learner corpora for this study regarding its creation discipline.

The latest version of the ICLE data (ICLEv3), which was released in 2020 as the culmination of the project, was used in this study. The main difference between the third version and others is that the third version has a web-based interface rather than the CD-ROMs provided with the first and second versions. This new interface also enables researchers to have faster and more flexible access to ICLE data. On the other hand, ICLEv3 includes 25 different L1 backgrounds. Three sub-

corpora: TICLE, NICLE, JPICLE as learner corpora and LOCNESS as a native learner corpus were chosen for this study. Moreover, the PTE Academic Collocation list was also used in this study to decide which *verb-noun* collocations would be analysed in this thesis.

The selection of the target corpora for this thesis is based on some strict criteria based on various variables such as corpus size, essay types, length of each essay, essay topics, the average age of the students, etc. Table 2 shows the comparability level of each corpus:

Table 2: Comparison Between the Target Corpora (ICLE and LOCNESS)

Variables	ICLE			LOCNESS	Comparability*
	TICLE	NICLE	JPICLE	Argumentative Literary-mixed	
Essay type	Argumentative	Argumentative	Argumentative	Argumentative Literary-mixed	++++
Number of Essays	280	317	366	322	++
Size	200,601 words	213,701 words	200,958 words	324,304 words	+++
Length of each essay (tokens)	+500	+300	+390	=/+ 500	++
The average age of students	22	23	20	17-22	++++
Level of study	University level (third to above years)	University level (first to the third year)	University level (first to fifth year)	University level (A-level students)	++
Topics	Divorce Abortion Sex Equality Nuclear Energy Suicide Capital Punishment Money Inventions Euthanasia Freedom of Press Human Rights Animal Testing University Degrees	Natural equality Feminism Professional military Science and technology Military army Crimes University degree Equality of men Dreaming and imagination Prison system Abortion	Language acquisition English as a second language American people	French intellectual tradition Loss of sovereignty for Britain Abortion Euthanasia Capital punishment Nuclear power Homosexuality Sex equality Crimes Feminists Divorce	+++
Time of completion	2016	2002	2005	1998	+

* +: low; ++: between low and average; +++: between average and high; ++++: high

Table 2 reveals the comparability criteria, graded as low, between low and average, between average and high, and high, for ICLE and LOCNESS; notably, the time of completion of the corpora exhibited low comparability. Comparability between low and average degrees was found for the length of each essay. Comparability between average and high was found for the size of the corpora that four corpora share a similar number of total words. Finally, high comparability was observed for essay types (argumentative) and the average age of students. The corpora were found

sufficiently comparable for this study. On the other hand, LOCNESS as a native corpus was also found comparable to ICLE because they both were composed of argumentative essays of university degree students with similar cognitive abilities.

All in all, each corpus is defined as a learner corpus sharing similar language levels, text and token size and same essay types. It should also be noted that the representation of three sub-corpora is wider in scope as they are selected from different geographies. One learner corpus was selected from Asia, Europe, and Eurasia to represent different linguistic backgrounds.

3.2.1. LOCNESS

Granger and Tyson (1996) state that the use of native corpus is an essential part of comparative corpus studies to enable a comparison of the data extracted from learner corpora. Thus, the Louvain Corpus of Native English Essays (LOCNESS) is used as a reference corpus in this study to control the non-native corpora (TICLE; NICLE; JPICLE). The corpus developed by Université Catholique de Louvain consists of 324,304 words collected from 322 argumentative and literary essays written by British and American university students and A-level essays by British students aged between 17 and 23. The average length of the essays is 500 words covering several topics such as transport, the parliamentary system, abortion and ethnic American literature. All the findings gathered from the three learner corpora were compared with the findings of the native corpus LOCNESS which is expected to represent the norms of native English speakers. The use of patterns and their overuse and underuse levels in ICLE data was compared with LOCNESS.

3.2.2. ICLEv3 (International Corpus of Learner English)

In 1990, Sylviane Granger introduced the International Corpus of Learner English aiming to provide non-native varieties of English (Granger et al., 2020: 1). The ICLE comprises written texts of argumentative essays which are written by EFL learners. There are three versions of the ICLE data. ICLEv1 published in 2002 contains 2.5 million words produced by EFL learners from 11 different mother-tongue backgrounds. On the other hand, ICLEv2 published in 2009 contains 3.7 million words written by EFL learners of 16 different L1 backgrounds. The latest version of ICLE, ICLEv3, has published in 2020 and the data consist of over 5.5 million words produced by learners from 25 different L1 backgrounds: Brazilian, Bulgarian, Chinese, Czech, Dutch, Finnish, French, German, Greek, Hungarian, Iranian, Italian, Japanese, Korean, Lithuanian, Macedonian, Norwegian, Pakistani, Polish, Russian, Serbian, Spanish, Swedish, Tswana and Turkish.

The ICLE includes 9529 essays totalling 5,766,522 words and the lengths of the essays ranges from 200,000 words to above. The learners consist of university level young adults in their twenties who are engaged in English as a Foreign Language (EFL) rather than English as a Second

Language (ESL). The proficiency of the learners, who are university undergraduates, ranges from higher intermediate to advanced. Moreover, most of the essays are argumentative, covering several topics such as ‘feminism’, ‘technology’, ‘money’ and ‘university ‘degrees’.

3.2.2.1. TICLE (Turkish International Corpus of Learner English)

The Turkish sub-corpus comprises 280 argumentative essays with a total word count of 200,601 collected from the Çukurova University, Mustafa Kemal University and Mersin University. The average length of the essays is 716.43. The average age of the learners in this sub-corpus is 19; in all, 81,43% of the learners are female and 18.57% are male. On the other hand, 20 randomly selected samples were rated on the basis of the CEFR category to decide the proficiency levels of each essay. Of the proficiency levels of all 20 sample essays, 16 were rated B2 and lower and 4 were rated C1 (Granger et al., 2020: 12).

TICLE was mainly selected as one of the learner corpora in this study for several reasons. I observed that Turkish English learners lack awareness of common word combinations such as collocations, phrasal verbs and idiomatic expressions. This is not only due to a lack of grammatical knowledge but also a lack of knowledge of the collocational nature of English. Foreign language learners suffer from these difficulties although the difficulty in the collocational nature of English might be challenging for a native speaker of English as well. However, teaching these word combinations enhances learners’ proficiency in the language (Lewis, 2000). Considering the situation in Turkey from the first point of view, TICLE was selected for this study to see the interlanguage production of Turkish EFL students. At the end of the study, a comparison of TICLE with other learner corpora and LOCNESS may help Turkish EFL learners to increase their proficiency in their written products (Granger et al., 2020).

3.2.2.2. NICLE (Norwegian International Corpus of Learner English)

The Norwegian sub-corpus contains 317 essays with a total word count of 213,701. Of the 317 essays, 312 are argumentative; 4 are literary essays and 1 is recorded as other. The essays are collected from 9 universities: the American College of Norway, Bergen College, Hedmark College, Oslo College, Ostfol College, University of Oslo, Agder College, Buskerud University and the University of Bergen. The average length of the essays is 674.14 and the average age of the learners is 18. Regarding the proficiency levels of all 20 essays, 8 were rated B2 and lower; 7 were rated C1 and 5 were rated C2. NICLE was selected because Norwegian learners showed more proficiency in English than Turkish learners (Granger et al., 2020: 12). On the other hand, the status of English in Norway was another reason for selecting NICLE. English is almost used as a second language in Norway as it became the core part of daily life as well as being taught in schools (Granger et al., 2020: 168).

3.2.2.3. JPICLE (Japanese International Corpus of Learner English)

The Japanese sub-corpus consists of 366 argumentative essays with a total word count of 200,958. The data was collected from 21 Japanese Universities: Waseda University, Showa Women's University, Keio University, Tokyo University of Foreign Studies, Tamagawa University, Aichi Shukutoku University, Dokkyo University, Fuji University, Hiroshima International University, Kooriyama Women's University, Kyoto University, Meiji University, Miyagi University of Education, Musashi University, Nihon University, Okayama University, Rikkyo University, Seijo University, Shinshu University, Shonann Institution of Technology and Tokai University. The average age of the learners is 20; in all, 72.95% are female and 26.50% are male. The average length of the essays is 549.07 whose proficiency level ranged from B2 (18 learners) to C1 (2 learners) according to the 20 sample essays. JPICLE was selected because Japanese learners showed similar proficiency to TICLE and NICLE. Also, the total number of words in both corpora has a similarity. (Granger et al., 2020: 12).

3.3. Data Analysis Tools

To analyse the corpus data in this thesis, two online tools were used: Sketch Engine and Pattern Dictionary of English Verbs (PDEV).

3.3.1. Sketch Engine

Sketch Engine was used to analyse the target corpora in this study. It is a free-available corpus tool providing researchers to study larger corpora through several functions such as word sketch, concordance, and n-grams. It serves as both a corpus software and a web service which covers corpus creation and management in addition to be a valuable tool for lexicographers, language teachers, universities, translators and technology companies (Kilgariff et al., 2014).

In this study, the word sketch (Figure 1) function, which shows a word's grammatical and collocational structures (Kilgariff et al., 2004), was used to generate a frequency list for collocations of *verb-noun*. *Word Sketch* function provides advanced search for filtering the results based on the parts of speech. It helps the examination of how words appear with a given word and are used grammatically; moreover, *Word Sketch* provide additional information about words' collocational tendencies and includes the potential collocations which were determined based on a set of rules. For example, first column in the figure 2 shows the modifiers of *allow* which provide more detail about the subject or object of a sentence. Second column, objects of "*allow*" displays the objects of the target verb that refers to the meanings, idioms and phrases occurred with the verb.

Figure 2: Screenshot of the Word Sketch Function in Sketch Engine

modifiers of "allow"	objects of "allow"	subjects of "allow"	prepositional phrases	pronominal objects of "allow"
then ... exotic surgery and then allowed to die slowly	boy ... boys are allowed	nature ... allowed by nature	"allow" by ...	her ... allow her
all ... allowed all	drug ... necessary to allow drugs	law ... allowed by the laws	"allow" in ...	them ... allow them to
only ... only allowed	same ... same be allowed	knowledge ... knowledge allows	"allow" for ...	him ... allowing him
not ... are not allowed to	euthenasia ... allow the euthenasia	husband ... husbands allow	"allow" over ...	us ... allows us
	euthanasia ... Euthanasia is only allowed	country ... country allow	"allow" with ...	it ... allow it
	everybody ... allowed , everybody	society ... allowed by some societies		
	pupil ... allowing pupils	teacher ... teachers allow		

Source: <http://www.sketchengine.eu/>, 2023.

Concordance is another function used in this study to show the keyword in context (KWIC). Concordance lines are also important in pattern studies as it provides observable usage of a pattern in these lines (Hunston, 2022). The usage of the target verbs and their patterns can be investigated in concordance lines. In order to provide reliable and valid data from the outcomes of the word sketches, concordance lines help the researcher to review the data in context and clean it if necessary. While using word sketch function, users may need additional information about how and where a specific word was used with other words.

3.3.2. The Pattern Dictionary of English Verbs (PDEV)

The Pattern Dictionary of English Verbs is the outcome of the Corpus Pattern Analysis method which was developed by Hanks (2013). It aims to annihilate the ambiguity between words and their meanings. The PDEV is a list of English verb patterns created by tagging verbs with their semantic types based on a custom ontology where each entry includes corpus-based CPA patterns and implicatures. The PDEV is used in this study as it provides more detailed descriptions of verb patterns matching meanings with verbs without grouping them into meaning groups because each verb has unique patterns.

3.4. Research Procedures

The main purpose of this thesis was to find out the similarities and differences of non-native learners' use of verb-noun patterns in their argumentative essays when they are compared to native use and each other. Initially, comparability of the target learners' corpora was done because it was important to study on the data which shared similar features such as text types, learner profile,

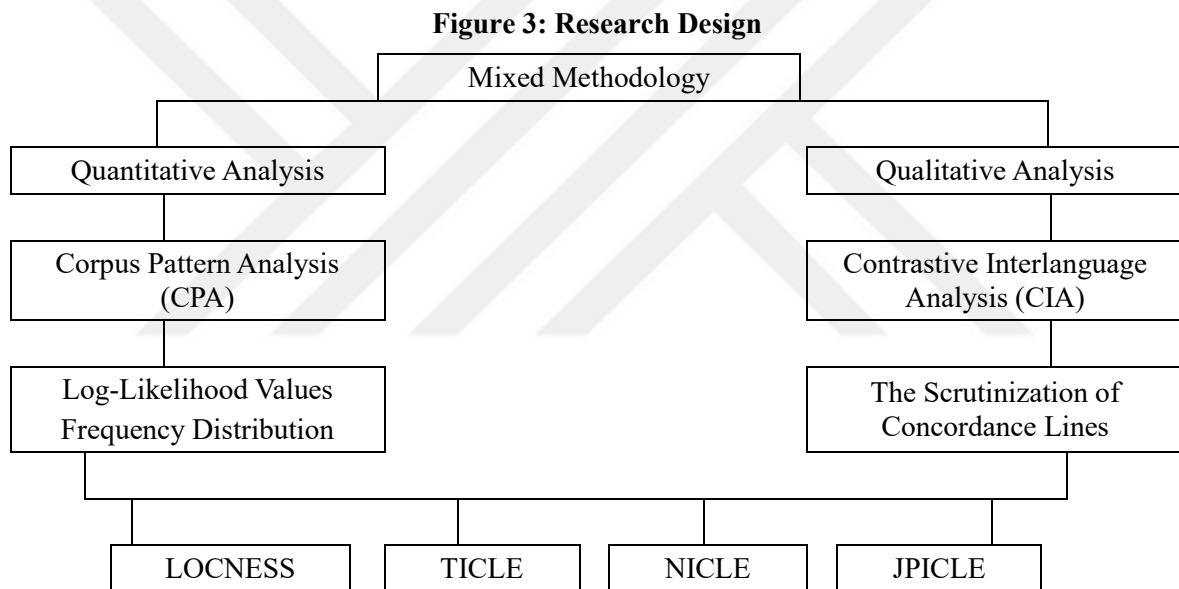
token and word size. On the other hand, the comparability of the reference data, which was LOCNESS, was another key point for the true representation and analysis of the data. In this study, the Turkish sub-corpus of ICLE (TICLE) was used as the researcher herself was a Turkish learner of English and to see the verb-noun use of Turkish learners to make further suggestions to improve their use if necessary. On the other hand, the Japanese sub-corpus of ICLE (JPICLE) was used due the common features shared by Japanese and Turkish learners such as the language family which was Altaic languages. The number of words, text types and learner variables in each corpus were also evaluated during the data selection process. The Norwegian sub-corpus of ICLE (NICLE) was chosen to compare Turkish and Japanese learner with a more proficient learner group which had closer relation with native speakers of English. The geographical diversity of the data was also showed how verb-noun patterns were used in different geographies. Lastly, three learner corpora (TICLE, NICLE and JPICLE) and a native reference corpus (LOCNESS) were uploaded to Sketch Engine for mixed analysis.

PTE Collocation List was used to choose which verbs would be studied as it included verb-noun collocations. The 63 verb-noun collocations were found from the list, and they were searched on Sketch Engine separately for the target corpora. It was one of the most important criteria for the selection of verbs was whether these verbs had studied in the PDEV or not. As a result, 10 verbs were found common in the PDEV, three learner corpora and a native corpus. However, the frequency of 3 verbs were low their patterns in the PDEV; thus, 7 verbs were ultimately chosen to analyse for this thesis. Initially, quantitative analyses were conducted to see the frequency of verbs in each corpus to compare. Then, statistical analyses were done to see if there was a statistical significance between the frequencies. According to the quantitative results, concordance lines were examined to provide qualitative description of the data. A detailed quantitative investigation was conducted to show the pattern distribution across the corpora. It showed the tendency of non-native learners towards the use of some patterns more.

This thesis was designed to maintain some necessary steps throughout the research. Figure 3 shows the steps which are planned to pursue. The overall aim of this thesis was to examine the verb-noun pattern use of non-native Turkish, Norwegian and Japanese learners in comparison with the use of native speakers of English. Mixed-method approach was applied to address the research questions. Mixed-method research is important in terms of providing deeper understanding of the results by combining the features of both qualitative and quantitative methods (Creswell, 2009: 204-205). It strengthens the thesis by using quantitative data from frequency and statistical analysis and qualitative data from concordance lines. Sequential embedded design was utilised that had two phases in this thesis. In sequential mixed research, the phases follow each other chronologically and one analysis follow the previous one (Teddlie & Tashakkori, 2008: 129). Firstly, frequency and statistical analysis of the verbs were conducted contrastively. The quantitative analyses were followed by examination of concordance lines to provide deeper understanding of the use of verb-

noun patterns. Another quantitative analysis was done to explain the diversity in concordance lines through log-likelihood and over-underuse of the verbs in specific patterns. Lastly, divergent usages of the verb-noun patterns were examined and matched their patterns in other sources which were *Grammar Patterns I: Verbs* book and BNC.

With quantitative analysis, corpus pattern analysis was mainly done. Moreover, the log-likelihood and frequency distribution of the verb-noun patterns across three learner corpora was calculated. On the other hand, Contrastive Interlanguage Analysis (CIA) was used for qualitative analysis. The concordance lines was listed and classified. For these analyses, three learner corpora (TICLE, NICLE, JPICLE) and a native corpus (LOCNESS) was used. To sum, this thesis was supposed to shed light on the over/underuse and erroneous or divergent use of verb-noun patterns. The findings of the analysis of the verbs are given in the following chapter.



CHAPTER FOUR

4. FINDINGS AND DISCUSSION

4.1. Data Analysis

4.1.1. Selection of Target Verbs

This thesis is a corpus-based contrastive study which compares the use of verb-noun patterns by non-native English learners and native speakers of English with a special focus to use the Pattern Dictionary of English Verbs (PDEV) to describe the data. The selection of the verbs to be examined in this study was based on some criteria. Initially, PTE Academic Collocation List (ACL) was used to decide which verbs would be used in the study. The list is created by Pearson (2018) and covers 2,469 most frequent written academic English collocations compiled from the Pearson International Corpus of Academic English (PICAE). 63 *verb-noun* (e.g. *allow access*) collocations were found from the list as a first step as these collocations are common in English. Table 3 shows the raw frequencies of verb-noun collocations in the target corpora. It is seen in the table 3 that the raw frequencies of the verbs are not the same across four corpora, and some verbs do not exist in them. For example, the verb *become* is mutually found 136 times in TICLE, 159 times in NICLE, 146 times in JPICLE and 171 times in LOCNESS. On the other hand, the verb *demonstrates* is found 3 times in TICLE and 11 times in LOCNESS while it does not occur in both NICLE and JPICLE. As a result, 21 verbs were found common in each corpus to be further investigated.

Table 3: Verbs and Their Raw Frequencies in Four Target Corpora (Verb-Noun Collocations)

LOCNESS		ICLE					
		TICLE		NICLE		JPICLE	
V	RF	V	RF	V	RF	V	RF
have	1281	have	1478	have	1227	have	1341
make	631	give	430	make	404	make	474
take	424	make	407	do	390	use	464
use	382	do	360	take	300	take	318
give	323	use	313	use	266	do	266
do	274	take	291	give	180	become	146
become	171	find	204	become	159	give	131
cause	160	become	136	find	113	find	105
show	157	cause	119	create	105	consider	98
find	153	create	85	commit	104	change	91
allow	147	provide	80	change	74	cause	64
increase	140	commit	75	consider	58	meet	55
create	133	develop	74	increase	49	develop	51
change	120	consider	69	develop	46	commit	50

Tablo 3: Continue

LOCNESS		ICLE					
		TICLE		NICLE		JPICLE	
V	RF	V	RF	V	RF	V	RF
provide	101	change	67	gain	36	show	48
accept	100	show	64	show	36	receive	38
consider	89	affect	61	meet	35	attend	34
reduce	79	accept	57	achieve	34	increase	32
receive	74	gain	40	discuss	34	create	26
present	71	reach	40	provide	30	carry	24
commit	70	apply	39	receive	29	follow	22
face	64	increase	36	reach	28	affect	21
discuss	55	require	32	experience	27	experience	21
raise	54	face	29	follow	25	accept	20
develop	52	discuss	26	require	25	establish	19
follow	49	achieve	25	allow	24	hold	19
require	48	establish	25	face	22	set	19
achieve	47	allow	24	cause	21	continue	18
affect	47	meet	23	raise	21	allow	17
experience	45	present	20	present	20	provide	15
hold	43	attend	18	set	19	reach	13
reach	42	perform	18	accept	18	convey	12
gain	41	carry	17	affect	18	discuss	12
establish	39	follow	16	attend	17	reduce	12
carry	37	continue	15	maintain	17	require	12
perform	35	draw	14	reduce	16	gain	11
set	35	hold	14	hold	15	raise	11
continue	33	experience	13	obtain	14	apply	9
adopt	30	adopt	12	carry	13	draw	9
obtain	30	reduce	12	draw	10	remain	9
draw	29	maintain	10	establish	9	achieve	8
meet	29	receive	10	perform	8	expand	8
maintain	27	add	9	add	7	face	7
describe	25	remain	9	seem	7	perform	7
add	23	seem	9	continue	6	collect	6
enhance	21	obtain	8	cover	6	describe	6
conduct	20	raise	7	describe	6	maintain	6
impose	20	conduct	6	remain	5	present	6
attend	19	cover	6	adopt	4	add	5
apply	17	impose	6	impose	4	conduct	5
remain	17	set	6	occur	4	occur	5
pose	15	occur	4	pose	4	adopt	4
assume	12	collect	3	apply	2	cover	4
expand	12	demonstrate	3	assume	2	impose	4
demonstrate	11	enhance	3	convey	2	identify	3
identify	10	describe	2	expand	2	seem	3
convey	9	expand	2	collect	1	obtain	2
cover	9	publish	2	conduct	1	pose	2
publish	8	assess	1	enhance	1	assume	1
collect	5	convey	1			enhance	1
occur	5	identify	1			publish	1
assess	2						

Further verb selection criteria were employed as the raw frequencies were not enough to investigate 21 verbs at the same time. To do this, the verbs were searched in PDEV to see if they already had been identified or not. Later, it was seen that these collocations had also been analysed for the PDEV project. The selection criteria were that every selected word from the PDEV project had at least 2 patterns. At the end of the search, 10 words were found common in PTE Academic

Collocation List and the PDEV project. Lastly, the frequency of the verbs was found in each of the target corpora (TICLE; NICLE; JPICLE) (Table 4 shows the frequencies of verbs in each source).

Table 4: Frequencies of Verbs in Each Source

VERBS	PDEV	PTE	LOCNESS	TICLE	NICLE	JPICLE
allow	8	1	147	24	25	17
accept	9	1	100	57	18	20
follow	20	3	49	16	25	22
affect	2	6	47	61	18	21
adopt	10	2	30	12	4	4
maintain	5	1	27	10	17	6
add	8	1	23	9	7	5
conduct	6	5	20	6	1	5
attend	3	1	19	18	17	34
apply	8	3	17	39	2	9

The columns in Table 4 show the pattern count for the PDEV project and the frequency of verbs in PDEV, PTE and target corpora. Objects of the “Verb” column in the Word Sketch function on Sketch Engine was used to extract the frequencies of verbs in four corpora. For example, table 4 shows that the verb “*follow*” has 20 patterns in PDEV and three occurrences in PTE Academic Collocation List. The highest frequency was found in the verb “*allow*”.

Further selection criteria were needed to be applied as a result of the less frequency of verbs in target corpora. For example, the verb *apply* has 8 patterns in PDEV but has 2 examples in NICLE. Similarly, verbs *adopt* and *conduct* has fewer examples in corpora than their pattern numbers in the PDEV. These verbs were eliminated because their frequencies were not enough to analyse and may not represent the exact usage. Ultimately, 7 verbs were chosen for further analyses: *accept*, *add*, *affect*, *allow*, *attend*, *follow*, and *maintain*.

4.1.2. Identifying Patterns

Patterns of the target words for this study were examined based on some steps described above.

1. All *verb+noun* collocations from the PTE Collocation List were gathered in a .docx format to be searched in target corpora and PDEV.
2. Each corpus was uploaded to Sketch Engine to generate frequency tables and concordance lines for the target verbs.
3. After uploading the ICLEv3 data, every *verb+noun* collocation gathered from PTE Collocation lists was searched to see their frequency in each corpus. Sketch Engine provided a search for all word forms at a time. Of the total number of 63 collocations, 21

were found in each corpus. There was also additional cleaning for the data to eliminate the less-frequent verbs. For example, the verb *identify* was searched and its raw frequency was either too low or non-existent in some of the corpora.

4. After generating concordance lines of the target verbs in KWIC format, the KWIC lines were copied from Sketch Engine and pasted to an Excel worksheet.
5. Concordance lines for every verb were manually studied and annotated according to the matching patterns already identified in the PDEV.
6. For example, 2 patterns were recorded in PDEV for the verb *affect*, and concordance lines from the TICLE are given in table 5.

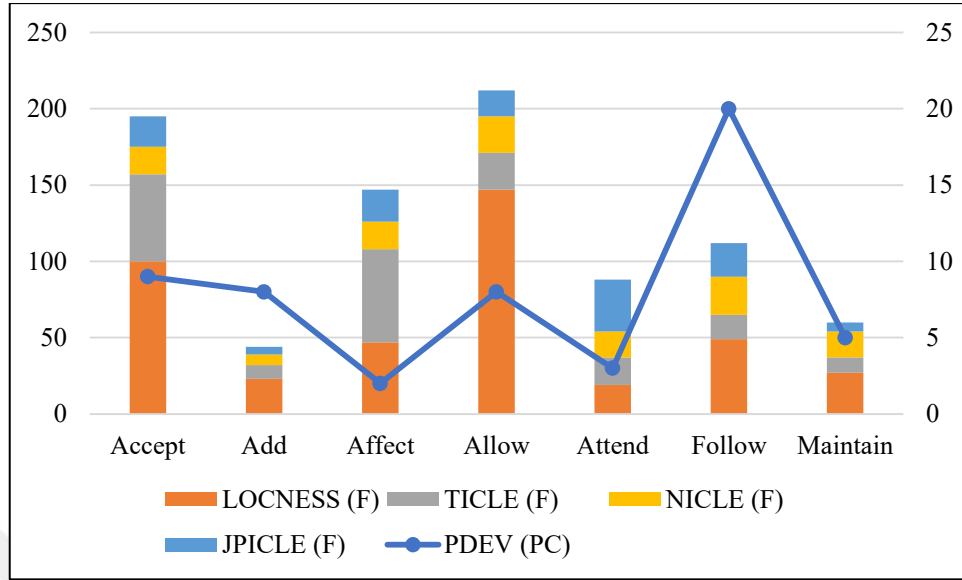
Table 5: PDEV Entry for Affect in TICLE

Pattern Number	Pattern and Implicature	TICLE (f)
1	Pattern: [[Eventuality 1 Entity1]] affect [[Human Animal Eventuality 2 Entity 2]] Implicature: [[Eventuality 1 Entity 1]] causes [[Human Animal Eventuality 2 Entity 2]] to change Example: It affects the marriage negatively, of course.	56
2	Pattern: [[Human]] affect [[Attitude Emotion]] Implicature: [[Human]] pretends to have [[Attitude]] or to feel [[Emotion]] Example: This lack of practice affects their success while performing their jobs.	5

In Table 5, column 1 indicates the pattern number of the verb *attend* in PDEV. Column 2 demonstrated the pattern and its implicature which is a dictionary definition to some extent. Lastly, 3rd column shows the frequency in the target corpus.

This chapter mainly focuses on the distribution of verb-noun patterns in native corpus LOCNESS and non-native sub-corpora TICLE, NICLE and JPICLE. Initially, it was important for verbs to be already investigated in the PDEV; as a result, 10 verbs were chosen out of 63 verb collocations in the PTE Academic Collocation list based on this selection criteria. Other 53 verbs were not investigated under the PDEV project. Previously, the frequency of 10 verbs was shown in Table 4; as a result, three verbs *adopt*, *apply*, and *conduct* were eliminated as they had less frequency in some corpora than their patterns in PDEV. For example, the verb *conduct* has 6 patterns in PDEV, but it has only one instance in NICLE and 5 instances in JPICLE. Thus, 7 verbs were ultimately chosen for further investigation as their frequencies were higher than their pattern numbers in the PDEV: *accept*, *add*, *affect*, *allow*, *attend*, *follow*, and *maintain*. Eventually, Figure 4 shows the distribution of target verbs with nouns in four corpora and their pattern number in the PDEV. Concordance lines of each verb were grouped according to the target corpora in an Excel (.xlsx) format to examine. Their pattern numbers were counted manually. The patterns and the implications of verbs are given in appendices (Appendix 1 for *accept*; Appendix 2 for *add*; Appendix 3 for *affect*; Appendix 4 for *allow*; Appendix 5 for *attend*; Appendix 6 for *follow*; Appendix 7 for *maintain*).

Figure 4: Target *Verb-n* Distribution



Concordantly, Log-likelihood and over-underuse values were calculated of these target verbs by using log-likelihood calculator provided online by Lancaster University to see their statistical significance in each corpus (Table 6). While calculating the LL value of the verbs, LOCNESS was used as second corpus to calculate the overuse and underuse values in learner corpora relative to native corpus.

Table 6: Log-Likelihood Values and Over/Under Representations

VERBS	TICLE		NICLE		JPICLE	
	Overuse Underuse	Log-Likelihood	Overuse Underuse	Log-Likelihood	Overuse Underuse	Log-Likelihood
accept	-	0.40 / $p < 0.05$	-	33.28 / $p < 0.0001$	-	26.67 / $p < 0.0001$
add	-	1.59 / $p < 0.05$	-	3.55 / $p < 0.01$	-	5.50 / $p < 0.01$
affect	+	13.83 / $p < 0.001$	-	4.01 / $p < 0.01$	-	1.60 / $p < 0.05$
allow	-	50.93 / $p < 0.0001$	-	53.82 / $p < 0.0001$	-	65.08 / $p < 0.0001$
attend	+	1.49 / $p < 0.05$	+	0.87 / $p < 0.05$	+	14.52 / $p < 0.001$
follow	-	5.83 / $p < 0.01$	-	1.05 / $p < 0.05$	-	1.62 / $p < 0.05$
maintain	-	2.25 / $p < 0.05$	-	0.02 / $p < 0.05$	-	6.26 / $p < 0.01$

The statistical investigation of target verbs shows that the verb *attend* was found over-used in target corpora and the verb *affect* was only found over-used in TICLE. As indicated in the table 6, the verb *allow* was under-used in the target corpora relative to LOCNESS with LL values of 50.93 in TICLE, 53.82 in NICLE, and 65.08 in JPICLE which is significant at the level of 99.99% ($p < 0.0001$). Similarly, the verb *accept* was also under-used in the target corpora with LL values of 33.28 in NICLE, 26.67 in JPICLE which is significant at the level of 99.99% ($p < 0.0001$); however, the LL value of the verb *accept* was found 0.40 which was significant at the level of 95% ($p < 0.05$). As it was beyond the scope of this thesis to investigate all the verbs, only the verbs *accept* and

allow were chosen for further investigation. Similarly, Alqarni (2019) studied the verb *allow* as it was found as the most noticeable verb in ICLE-Swedish with the LL value of 44.51 and it was underused in ICLE-Swedish when it compared to LOCNESS. Following sections will focus on the detailed investigation of target verbs in each corpus.

4.2. Identifying *accept-n* Patterns in Target Corpora

This section focuses on the comparison between TICLE, NICLE, JPICLE and LOCNESS on the use of *accept-n* patterns. The raw frequency of *accept-n* usages in target corpora was found 195 (TICLE $f= 57$; NICLE $f=18$; JPICLE $f= 20$; LOCNESS $f= 100$). These 195 lines were examined individually and matched the patterns already exist in the PDEV. Moreover, their log-likelihood (LL) scores and over/underuse values were calculated with the online log-likelihood calculator provided by Lancaster University to decide whether the frequency differences in three corpora have a statistical significance or not. Table 7 shows the raw frequencies and LL scores of *accept-n* in target corpora.

Table 7: Statistical Aanalysis of *accept-n* Patterns

Corpora	Frequency	Log-likelihood Scores	Overuse / Underuse
TICLE	57	0.40 $p<0.05$	-
NICLE	18	33.28 $p<0.0001$	-
JPICLE	20	26.67 $p<0.0001$	-

According to Table 7, *accept-n* patterns have 57 occurrences in TICLE, 18 in NICLE and 20 in JPICLE. As indicated in Table 7, *accept-n* patterns are under-used in each corpus relative to the native corpus LOCNESS with a log-likelihood score of 0.40 for TICLE; 33.28 for NICLE; and 26.67 for JPICLE. It is seen that *accept-n* patterns are less represented in NICLE and JPICLE. Although the target verb-noun pattern is under-used in each corpus, it is also notable that *accept-n* use in TICLE shows the highest significance with the LL score of 0.40 which is significant at the level of 95% ($p<0.05$).

The *accept-n* patterns found in the corpora were matched to the patterns in PDEV. There were 9 patterns of the verb *accept* in PDEV; however, only 6 patterns were matched to the instances of the target corpora. Table 8 shows the PDEV entries to see the similar and different use of the target verb in the corpora.

Table 8: PDEV Entries of Accept-N Patterns

Pattern Number	Pattern and Implicature	LOCNESS ϕ
1	Pattern: [[Human Institution]] accept [[Proposition Concept Eventuality]]	82
	Implicature: [[Human Institution]] agrees that [[Proposition Concept Eventuality]] is correct and does not need to be contested	
	Example: If a human being has a brain and hearth, he can't accept this punishment . (TICLE)	
	Example: He accused them of accepting the doctrines of their religion without reflection. (NICLE)	
	Example: The fact that people in the south of Thailand refuse to accept the guidelines . (JPICLE)	
2	Pattern: [[Human Institution]] accept	3
	Implicature: [[Human Institution]] agrees that {that [CLAUSE]} is true or correct	
	Example: Shortly we cannot accept the idea that there is a certain sex equality or sex discrimination. (TICLE)	
	Example: He did not want to accept the fact that he was increasingly failing as a salesman. (NICLE)	
	Example: If you become the entertainers, you should accept the fact that your life is always looked by others or media. (JPICLE)	
5	Pattern: [[Human]] accept [[Offer]] {offer proposal recommendation suggestion advice}	1
	Implicature: [[Human 1]] agrees to act on the [[Offer]] of [[Human 2]]	
	Example: Accepting this viewpoint , you accept that someone should be able to decide to die. (TICLE)	
	Example: I would guess that... not worrying whether or not the "big client" will accept their offer next week... (NICLE)	
	Example: No doubt a majority opinion is regarded as important in the society where people accept a majority decision . (JPICLE)	
8	Pattern: [[Human 1 Institution 1]] accept [[Human 2 Institution 2]] [as [[Human_Role Institution_Role]]]	5
	Implicature: [[Human 1 Institution 1]] acknowledges that [[Human 2 Institution 2]] is suitable for [[Human_Role Institution_Role]]	
	Example: The men should accept women as an important power in the society and family. (TICLE)	
	Example: Even though the first blacks came centuries ago, blacks were not accepted as equal citizens. (NICLE)	
9	Pattern: [[Human_Group 1 Human 1]] accept [[Human 2 Human_Group 2]]	-
	Implicature: [[Human_Group 1 Human 1]] agrees to allow [[Human 2 Human_Group 2]] to become a member of or be respected by [[Human_Group 1]]	
	Example: But shortly, they have started to accept students by entrance exam score alone. (TICLE)	
	Example: We cannot touch precious a social internationalization or information-oriented time and superficial even accept foreign cultures . (JPICLE)	

Table 8 shows the 6 patterns of *accept-n* found in the PDEV. Column 1 indicates the pattern number in the PDEV for the target verb. Column 2 shows the pattern, its description (implicature), and examples from each corpus. Column 3 indicates the pattern frequencies that occurred in LOCNESS. There were no matches of patterns 4, 6, and 7 (Appendix 1) in the target corpora. On the other hand, some patterns were not found mutual in some corpus. For example, patterns 1, 2 and 5 were found in each corpus; pattern 8 was only found in both TICLE and NICLE. Another finding showed that it was not found any matches of pattern 9 in LOCNESS while TICLE and JPICLE had at least one instance.

Table 9: Log-likelihood and Over/Under Representations Of Patterns For Accept-N

Pattern Number	1	2	5	8	9
TICLE	2.13 / -	2.25 / +	0.16 / +	10.98 / +	-
NICLE	0.47 / -	0.26 / +	1.32 / +	2.35 / +	-
JPICLE	1.10 / -	1.53 / +	9.69 / +	-	-

The table 9 above gives statistically detailed information about the patterns in each corpus. The LL value of pattern 9 was not calculated because there were not instances in LOCNESS. It is notable from Table 9 that pattern 8 has been overused in TICLE with the LL value of 10.98 which is significant at 99.99%. Moreover, pattern 5 is also overused in JPICLE with the LL value of 9.69 which is significant at 99.9%. An investigation of concordance lines for pattern 8 in TICLE showed 14 occurrences. According to pattern 8 whose description was given in Table 9, a human or institution is followed by a human or institution as a role. An example from LOCNESS and TICLE is given below.

4.1. In the novel, the young Candide **accepted Optimism as** an unshakeable theory when he was taught it by his tutor. (LOCNESS)

4.2. **Women are also accepted as** the one who do the housework, cooking, taking care of the children, shopping, etc. (TICLE)

As shown in Table 9 *Accept* in pattern 8 is followed by a human or institution as a role. There was needed syntactic alternation for example 4.2. to analyse the sample. Hanks (2013: 188) indicated that “if the subject of a passive verb is the same as the direct object of the same word as an active verb, it is generally not necessary to propose separate patterns for the active and passive uses of that verb”. Example 4.1. was active and 4.2. was passive; however, the relationship between the verb *accept* and both examples is the same.

It is apparent from the statistical analyses given in this section that the verb *accept* occurs in four patterns with a noun in target corpora. Pattern 9 was used in TICLE and JPICLE and not used in NICLE and LOCNESS; thus, the log-likelihood values of pattern 9 across corpora could not be calculated as the reference source did not include the same data.

4.3. Identifying *Allow-n* Patterns in Target Corpora

The detailed investigation of *Allow-n* patterns in three target corpora is given in this section. As a result of the quantitative investigation of target verbs showed in the table 6, *Allow* was found as the most noticeable verb which was significant at 99.99% ($p < 0.0001$). It had also 8 patterns in PDEV and the highest frequency in the target corpora which made it a good choice to investigate. The raw frequency of *Allow-n* usages in target corpora was found 212 (TICLE $f= 24$; NICLE $f=24$;

JPICLE $f= 17$; LOCNESS $f= 147$). 212 lines were manually analysed and matched to the patterns in the PDEV. Table 10 shows the log-likelihood and over-underuse values of the verb *allow* in each corpus.

Table 10: Statistical Analysis of *allow-n* Patterns

Allow-n	Frequency	Log-likelihood Score	Overuse / Underuse
TICLE	24	50.93 / $p < 0.0001$	-
NICLE	24	53.82 / $p < 0.0001$	-
JPICLE	17	65.08 / $p < 0.0001$	-

The LL values indicated that *allow-n* patterns were underused in target corpora relative to LOCNESS. Their LL values (50.93 in TICLE; 53.82 in NICLE; 65.08 in JPICLE) were found significant at 99.99% ($p < 0.0001$). As it was mentioned before, *allow* has 8 patterns in the PDEV; however, there were 4 patterns matched with the instances in the target corpora. Table 11 presents the matched patterns of *allow-n* in the PDEV.

Table 11: PDEV Entries of *allow-n* Patterns

Pattern Number	Pattern and Implicature	LOCNESS ϕ
1	Pattern: [[Human 1 Institution 1 Eventuality]] allow [[Human 2 Institution 2 Self]] [to+INF]	97
	Implicature: [[Human 1 Institution 1 Eventuality]] gives [[Human 2 Institution 2 Self]] the opportunity, time, or permission {to/INF [V]}	
	Example: You should not allow the money to rule you, you should rule it according to your desires. (TICLE)	
	Example: These games allow people to use their imagination at a high level. (NICLE)	
	Example: The US should allow local people to establish their own government. (JPICLE)	
2	Pattern: [[Human 1 Institution 1 Eventuality]] allow [[Institution 2 Self Privilege Human 2]]	8
	Implicature: [[Human 1 Institution 1 Eventuality]] gives [[Human 2 Institution 2 Self]] permission or the opportunity to have [[Privilege]]	
	Example: the risk of euthanasia is ... allowing individuals the freedom to say "here is as far as I go". (TICLE)	
	Example: I think that they should be allowed more play time , where they can use their imagination more freely. (NICLE)	
	Example: Such education is not education that does not allow students opportunities to learn real life. (JPICLE)	
3	Pattern: [[Eventuality Human Institution]] allow [[Physical_Object]] {to be V-ED} [to+INF]	7
	Implicature: [[Human Institution Eventuality]] causes or provides the opportunity for [[Physical_Object]] {to be V-ED}	
	Example: On the other hand, huge public support that animal testing is necessary to allow drugs to be safely tested on and used by humans. (TICLE)	
	Example: Why can't children be allowed to be bored sometimes... (NICLE)	
4	Pattern: [[Human Eventuality 1 Institution]] allow [[Eventuality 2]]	31
	Implicature: [[Human Institution Eventuality 1]] gives the opportunity, time, or permission for [[Eventuality 2]] to occur	
	Example: Remember that just standing idle when you could prevent death is a decision to allow suicide . (TICLE)	
	Example: Even so, we allow a certain amount of censorship in our media. (NICLE)	
	Example: Japan administration obey the OECD's logic and doesn't allow the necessity of "Indication" . (JPICLE)	

There are 8 patterns in the PDEV for the verb *allow*; however, it was only 4 patterns found for the verb in the target corpora in this thesis. There were no instances for the pattern numbers 5, 6, 7 and 8 (Appendix 4). Column 1 shows the pattern number in the PDEV, Column 2 describes the pattern and its implicature with examples from the target corpora. It should be noted that there were no instances of pattern 3 in JPICLE found. A more detailed investigation was conducted to reveal the frequency distribution among patterns in each corpus. Table 12 summarises the statistical values of patterns.

Table 12: Log-likelihood and Over/Under Representations of Patterns for *allow-n*

Pattern Number	1	2	3	4
TICLE	4.19 / -	0.07 / -	1.68 / +	5.58 / +
NICLE	0.00 / +	0.27 / +	0.02 / -	0.00 / -
JPICLE	0.12 / -	0.01 / +	-	1.19 / +

The quantitative analysis of the target verb according to its patterns across target corpora revealed that the frequencies in corpora are less than those in LOCNESS. Pattern 4 was over-used in TICLE with the LL value of 5.58 which is significant at 99% ($p < 0.01$). Another finding showed that Pattern 1 was under-used in TICLE with the LL value of 4.19 which is significant at 99% ($p < 0.01$). A total of 24 instances were found for *allow-n* and an investigation for pattern 4 in TICLE showed 12 matches. On the other hand, of the 147 instances for the target verb in LOCNESS, 31 of them were matched with pattern 4 in the PDEV. The following example shows the use of *allow-n* pattern 4 in both TICLE and LOCNESS.

4.3. The power granted under article twelve to dissolve parliament would, according to Michel Debré, only **allow a short exchange** with the public. (LOCNESS)

In pattern 4, the implicature is *Human or Institution or Eventuality 1 gives the opportunity, time, or permission for Eventuality 2 to occur*. The description of example 4.3. is *Eventuality 1 allows Eventuality 2*.

4.4. Thus, not all the societies want to **allow the euthenasia** because of their cultural and moral norms I'm agree with the application of the euthenasia. (TICLE)

4.5. This is also a religious issue for many people, because of the fact that most religions don't **allow abortion**. (NICLE)

4.6. In big cities most apartments do not **allow pets**. (JPICLE)

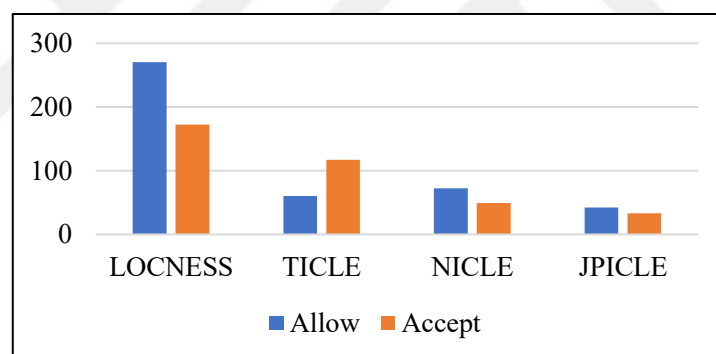
Moreover, a linguistic phenomenon “*metonymy*” was observed in examples 4.4., 4.5., and 4.6. that the term is defined as the use of a salient entity to refer to another entity (Littlemore, 2015). Regarding the PDEV, it is defined as ‘a semantic type alternation’ referring to the focus on

an entity which is affected by another one (Hanks, 2013). In example 4.4., ‘*the societies*’ alternates with the government which allows euthanasia. The term ‘*religions*’ refers to in example 4.5. alternates with the. In example 4.6., ‘*apartments*’ refers to the semantic type [[Social Location]] which alternates with the semantic type [[Human Institution]] referring to the building management or building residents. As Hanks (2018: 177) indicated that “there is a regular alternation in different kinds of contexts between lexical sets that have different semantic types”. Further analysis was required to explain the learner corpora and to determine the attitudes of Turkish, Norwegian, and Japanese learners of English towards employing verb patterns in their writings. Consequently, error analysis was run to the raw usages of *accept* and *allow* in the following section.

4.4. Divergency of *Accept* and *Allow*

In this section, the overall evaluation of divergent use of *accept* and *allow* occurred in three corpora was demonstrated. Firstly, Figure 5 shows the raw frequencies of the verbs *accept* and *allow* in the target corpora.

Figure 5: *accept* and *allow* Distribution across Corpora



It is clearly seen from the figure 5 that JPICLE -as a learner corpus- has the lowest frequency values for *accept* and *allow* while TICLE and NICLE share the similar frequencies. However, the use of *accept* in TICLE is found noticeably high among the corpora. More detailed investigation was made to identify divergent use of verb patterns by Turkish, Norwegian, and Japanese learners of English. Firstly, concordance lines of *accept* and *allow* were matched with their patterns in PDEV as presented in previous section. Later, some concordance lines could not be matched with any patterns in PDEV, and it was decided to analyse them in detail. *Grammar Patterns 1: Verbs* and *British National Corpus (BNC)* were used to decide whether these concordance lines were divergent or not. Some patterns not found in PDEV were detected and the examples were searched in BNC or their patterns were searched in *Grammar Patterns 1: Verbs* book to call them as divergent.

4.4.1. Focus on *Accept* in Target Corpora

CPA method was applied to describe the corpora whose concordance lines were examined and matched with the patterns in PDEV. 9 patterns were recorded in PDEV for *accept* based on the 250 instances in BNC50 data. On the other hand, Table 13 shows the raw frequency of *accept* in target corpora. It is seen that the RF is 172 in LOCNESS.

Table 13: Raw Frequency and LL Values of *accept*

<i>Accept-v</i>	Raw Frequency		
LOCNESS	172	Log-likelihood Score	Overuse / Underuse
TICLE	117	0.36/p < 0.05	+
NICLE	49	30.22/ p < 0.0001	-
JPICLE	33	48.25/ p < 0.0001	-

In all, 371 concordance lines of *accept* were examined and only 25 did not match any of the 9 patterns in PDEV. TICLE has 117 occurrences for *accept* and 9 of them were found as possible divergent instances. Figure 6 shows the concordance lines in TICLE, which did not match any pattern in PDEV.

Figure 6: Concordance Lines of *accept* in TICLE which Did not Match Patterns in the PDEV

1	to find an institution which will	<i>accept</i>	them to internship programs.
2	history, many cultures have	<i>accepted</i>	euthanasia as a normal act.
3	are important issues which are	<i>accepted</i>	as handicap and discussed in society.
4	well as objections to it and it is	<i>accepted</i>	as a big discussing point in religion, medicine
5	from this perspective I could	<i>accept</i>	euthanasia as a remedy which stops patient's
6	They are	<i>accepted</i>	like a servant.
7	the world, nuclear power is	<i>accepted</i>	as the only useful power for all the benefits of
8	the university degree will be	<i>accepted</i>	a real value for the real world if the lessons and
9	She does not	<i>accept</i>	that as a normal situation of her body,

These instances did not match any pattern; however, further investigation was done to decide whether they were divergent or not. The instances were examined based on the patterns in *Grammar Patterns I: Verbs* book. As a result, 7 of 9 instances were matched their patterns found in the *Grammar Patterns* but not in PDEV. The classification of their patterns in *Grammar Patterns* book is given in Table 14.

Table 14: Pattern Classification of *accept* in TICLE Instances Not Matched in PDEV

Classification	Pattern	Example	BNC (f)
Complex patterns with prepositions and adverbs	• V n to n	<i>accept</i> them to internship programs	39
	• V n as n	<i>accept</i> euthanasia as a remedy	57
	• V n for n	<i>accept</i> a real value for the real world	200

Since verb pattern in Table 14 (*accept* followed by noun group and like) was not found in *Grammar Patterns 1: Verbs*, the pattern was searched in BNC. The search was done by using collocates function in BNC by entering *accept* NOUN to verb/phrase row and *like* to collocate row. This verb pattern was used 39 times in BNC which showed low frequency in such a big data. Thus, this pattern was left for further examination to decide whether it was divergent or not.

On the other hand, NICLE has 49 instances of *accept* and only three of them did not match any pattern in PDEV. Figure 7 displays the instances of concordance lines in NICLE that did not accord to patterns in PDEV.

Figure 7: Concordance Lines of *accept* in NICLE which Did not Match Patterns in PDEV

1	and even seem to be	<i>accepted</i>	as necessary
2	when they are not	<i>accepted</i>	for who they are?
3	man who was unlikely to	<i>accept</i>	help from others.

No pattern was found to match the pattern in PDEV, they were scrutinized to indicate whether they were divergent or not. The investigation of *accept* focused on the patterns in *Grammar Patterns*. Subsequently, all instances were found to have patterns in the *Grammar Patterns* book and instances in BNC. The classification of these patterns and their BNC frequencies are depicted in Table 16.

Table 15: Pattern Classification of *accept* in NICLE Instances Not Matched in PDEV

Classification	Pattern	Example	BNC (f)
Complex patterns with prepositions and adverbs	• V n as adj	<i>accept</i> discrimination as necessary	0
	• V n for n	<i>accept</i> them for who they are	113
	• V n from n	<i>accept</i> help from others	88

In the first classification (V n as adj), a noun group, a prepositional phrase with *as* and an adjective follow the verb by indicating a specific quality or description of the object of the sentence. Then, for the use of “V n for n” pattern, the verb of the example (*accept*) is followed by a noun and a prepositional phrase with *for* which describes the noun. Lastly, the pattern “V n from n” refers acquiring something from someone or somewhere that the example means being willing to receive support *from* someone.

5 instances of *accept* in JPICL did not match any pattern in the PDEV. According to the investigation of these instances, four instances had their patterns in *Grammar Pattern 1: Verbs*, and one instance could not be defined. Figure 8 shows the occurrences of concordance lines in JPICL that did not match the patterns in PDEV.

Figure 8: Concordance Lines of *accept* in JPICL which Did not Match Patterns in PDEV

1	I should think that teachers	<i>accept</i>	using English to teach other subject
2	nuclear energy will be	<i>accepted</i>	to society, if it is used peacefully
3	more houses will	<i>accept</i>	having pets at home
4	So, if we	<i>accept</i>	benefit from other countries
5	at least for moderns, to	<i>accept</i>	sudi a condition.

The first four instances matched patterns in *Grammar Patterns* book as complex patterns and complex patterns with prepositions and adverbs. However, the fifth occurrence could not be defined because the term “*sudi*” could not find in BNC, and it was marked as an error. On the other hand, table 16 shows the classification of the patterns for the remaining instances in JPICL with their BNC frequencies.

Table 16: Pattern Classification of *accept* in JPICL Instances not Matched in PDEV

Classification	Pattern	Example	BNC (f)
Complex Patterns	V n <i>to-inf</i>	<i>accept</i> using English to teach	44
Complex patterns with prepositions and adverbs	• V n <i>to</i> n	<i>accept</i> nuclear energy to society	30
	• V n <i>at</i> n	<i>accept</i> having pets at home	8
	• V n <i>from</i> n	<i>accept</i> benefit from other countries	88

The patterns of *accept* were classified in *Grammar Patterns* book as complex patterns and complex patterns with prepositions and adverbs. In its first pattern, *accept* was followed by noun group and infinitive (V n *to-inf*) as a complex pattern. The example (*accept using English to teach*) showed the tendency of approving a noun group for a specific purpose, and *using English* was used as a noun group. Three patterns for *accept* in JPICL were found as complex patterns with prepositions and adverbs. First pattern of this classification (V n *to* n) showed that “*to*” was used as a preposition rather than infinitive, and it contributed the passive meaning to the sentence. It was found that “*to*” in this instance shared the similar meaning with “*by*” in a passive form to indicate who would accept the eventuality. Other prepositions *at* and *from* was used to refer the place of the eventuality.

4.4.2. Focus on *Allow* in Target Corpora

CPA method was used to provide a comprehensive interpretation of the corpora. Meticulous investigation of *allow* was conducted on the concordance lines; as a result, the concordance lines

were matched their patterns in PDEV where 8 patterns were recorded for *allow*. Table 17 shows the raw frequency and log-likelihood values of *allow* in each corpus.

Table 17: Raw Frequency and LL Values of *allow*

<i>Allow-v</i>	Raw Frequency		
LOCNESS	270	Log-likelihood Score	Overuse / Underuse
TICLE	60	65.58 / p < 0.0001	-
NICLE	72	53.39 / p < 0.0001	-
JPICLE	42	94.38 / p < 0.0001	-

According to the table 17, the raw frequency of *allow* in reference corpus LOCNESS was 270. Log-likelihood values of *allow* in target corpora was found 65.58 in TICLE, 53.39 in NICLE and 94.38 in JPICLE which indicated the association between the observed data and LOCNESS data. The 99.99% significance level and p-value emphasised the significance of the observed data. Overall, the data found in the three corpora showed significant association; moreover, the underuse of the data demonstrated that the corpora might be further investigated.

In total, 444 concordance lines of *allow* were examined, with only 8 of them did not match any of the 8 patterns present in the PDEV. The raw frequency of *allow* in TICLE was 60 and 3 instances were identified as potential divergent patterns. These instances are given in the following figure 9 which demonstrates the concordance lines in TICLE that did not match any pattern in the PDEV.

Figure 9: Concordance Lines of *allow* in TICLE which Did not Match Patterns in PDEV

1	Why shouldn't the same be	<i>allow</i>	with humans, if pain reaches a level
2	While some country	<i>allow</i>	it many country in the world
3	If they don't	<i>allow</i>	to lose its value, it will teach them everything

As indicated in the Table 17, *allow* was underused in TICLE with an LL value of 65.58 which was significant at 99.99%. 3 of the 60 *allow* instances in TICLE were found possible divergent usages. Thus, they were further examined by checking their patterns in *Grammar Pattern Book 1: Verbs* and *BNC*. The table 18 shows the classification of these patterns with their BNC frequencies.

Table 18: Pattern Classification of *allow* in TICLE Instances Not Matched in PDEV

Classification	Pattern	Example	BNC (f)
Simple Patterns	V <i>to-inf</i>	<i>allow</i> to lose	153
Simple Patterns with Prepositions and Adverbs	V <i>with n</i>	<i>allow</i> with humans	0
Complex Patterns with Prepositions and Adverbs	V n <i>in n</i>	<i>allow</i> it many country in the world	35

Each instances matched patterns in *Grammar Patterns* book within three classification types which are *simple patterns* (v *to-inf*), *simple patterns with prepositions and adverbs* (v *with n*) and *complex patterns with prepositions and adverbs* (v n *in n*). Simple verb patterns in this book have meaning groups such as *begin, appear, try, manage, fail, regret to say, hasten, chance, and tend*. TICLE example with simple pattern “allow to lose” matched with begin group which primarily focus on the initiation and maintenance of an action. Moreover, second pattern group consists of *allow* which was followed by a prepositional phrase comprising “*with*” and a noun group (allow with humans). Lastly, the verb *allow* is followed by a noun group and a prepositional phrase with “*in*” followed by a noun group (allow it many country in the world).

Further investigation of *allow* was conducted on NICLE which involves 72 instances of *allow*. 2 of the 72 instances did not match any pattern in the PDEV. Figure 10 demonstrates the possible divergent instances.

Figure 10: Concordance Lines of *allow* in NICLE which Did not Match Patterns in PDEV

1	more than ever, it should be	<i>allowed</i>	to sit back and relax with a good book
2	if they only were	<i>allowed</i>	to be airplanes once in a while

The verb *allow* was underused in NICLE with an LL value of 53.39 which was significant at 99.99%. As two instances did not match any pattern in the PDEV, further investigation was undertaken by checking their patterns in *Grammar Patterns 1: Verbs* book and their use in BNC. As a result, the classification of these instances is given in Table 19.

Table 19: Pattern Classification of *allow* in NICLE Instances Not Matched in PDEV

Classification	Pattern	Example	BNC (f)
Simple Patterns	V <i>to-inf</i>	<i>allowed</i> to sit back <i>allowed</i> to be airplanes	153

Both instances matched patterns in *Grammar Patterns* book as simple pattern. The verb *allow* is followed by “*to-infinitive*”. In second instance (allowed to be airplanes), the verb (be) is followed by an object (airplanes).

Lastly, the same procedure was done with instances in JPICLE. 3 of 42 instances did not match any patterns in the PDEV and were given in Figure 11.

Figure 11: Concordance Lines of *allow* in JPICLE which Did not Match Patterns in PDEV

1	there are many reasons not to	<i>allow</i>	to give up
2	I wonder it should be	<i>allowed</i>	to intrude when one has baby.
3	name without any procedures	<i>allows</i>	using woman's body as a tool for childbearing

The verb *allow* was underused in JPICLE with a highest LL value of 94.38 which was significant at 99.99%. As a result of the detailed investigation, three instances matched patterns *Grammar Patterns* book. The classification of these patterns is given in Table 20. The instances were classified as *simple pattern (v to-infinitive)* and *complex pattern with preposition (v n as n)*.

Table 20: Pattern Classification of *allow* in JPICLE Instances Not Matched in PDEV

Classification	Pattern	Example	BNC (f)
Simple Patterns	V to-inf	<i>allow</i> to give up	153
Complex Patterns with Prepositions and Adverbs	V n as n	<i>allow</i> using woman's body <i>as</i> a tool	4

4.5. Discussion of the Research Questions

This thesis employed Corpus Pattern Analysis by comparing three learner corpora (TICLE, NICLE and JPICLE) with a native reference corpus (LOCNESS) to identify the use of verb-noun patterns. The first half of this chapter presented the findings of qualitative and quantitative analysis. This section focuses on the discussion of the research questions presented in the introduction chapter in this thesis. The research questions which pave the way of this thesis are as follows:

1. What are the similarities and differences between verb-noun patterns used by learners in TICLE, NICLE, JPICLE and the LOCNESS corpora?
2. What are the probable divergent patterns in the use of verb patterns in the learner corpora?
3. How can CPA be used to describe verb-noun patterns in learner corpora with different L1 background?
4. To what extent is the use of PDEV for the description and identification of target divergent patterns possible in learner corpus analysis?

4.5.1. Discussion of the First Research Question

This thesis is a contrastive corpus study which aimed to investigate the use of verb-noun patterns among learner corpora; therefore, it asked the question whether there were similarities and differences between three learner corpora when LOCNESS data was taken as a reference. As a result of several selection criteria, 7 verbs were chosen for detailed examination. It was crucial to select verbs that had higher frequency than their pattern numbers in the PDEV to decide their pattern use based on their pattern entries in the PDEV. The statistical investigation which was

given in Table 6 showed that the verbs *accept* and *allow* were underused in each learner corpus with a significance level of 99.99%. The investigation of both verbs for their patterns revealed that not all corpora had instances for each pattern in the PDEV. For example, the verb *accept* has nine patterns in the PDEV; however, the verb *accept* was used in four patterns in target corpora. Moreover, learner corpora had frequency in Pattern 9 (Pattern: *[[Human_Group 1 | Human 1]] accept [[Human 2 | Human_Group 2]]*) but the log-likelihood could not be calculated because LOCNESS as a reference corpus did not have any instances for this pattern. It was seen in TICLE data that learners used pattern 8 (see Appendix 1) instead of pattern 9; this means that Turkish learners tend to use “as” after the semantic type “*Human, Human Group or Institution*” follows the verb *accept*. These findings broadly corroborate the ideas of Sinclair (1991: 53) about the relationship between sense and meaning which indicates the meaning of a specific word is related to a specific pattern. Another possible reason for the similarity between TICLE / JPICLE and NICLE / LOCNESS might be their proficiency levels. As ICLEv3 Manuel (2020) indicated that English is widely used in Norway in addition to Norwegian language which indicates Norwegian learners of English show more proficient learner profile than Turkish and Japanese learners. Additionally, NICLE and LOCNESS might be expected to show similar pattern use in their analysis. Table 21 shows the verb-noun patterns of *accept* in learner and native corpora.

Table 21: Verb-Noun Patterns of *accept* in Target Corpora

Pattern Number	1	2	3	4	5	8	9
LOCNESS	82	3	4	2	0	5	0
TICLE	36	5	1	0	1	14	1
NICLE	12	1	0	0	1	1	0
JPICLE	12	2	0	0	4	0	2

Moreover, Table 22 presents the verb-noun patterns of *allow* in learner and native corpora.

Table 22: Verb-Noun Patterns of *allow* in Target Corpora

Pattern Number	1	2	3	4	5	6
LOCNESS	97	8	7	31	1	3
TICLE	8	1	3	12	0	0
NICLE	16	2	1	5	0	0
JPICLE	10	1	0	6	0	0

Of the 9 patterns of *accept* in the PDEV, 5 patterns were used in LOCNESS (1, 2, 3, 4, 8); 6 patterns were used in TICLE (1, 2, 3, 5, 8, 9); 4 patterns were used in NICLE (1, 2, 5, 8) and JPICLE (1, 2, 5, 9). On the other hand, of the 8 patterns of *allow* in the PDEV, 6 patterns were used in LOCNESS (1, 2, 3, 4, 5, 6); 4 patterns were used in TICLE (1, 2, 3, 4) and NICLE (1, 2, 3, 4); and 3 patterns were used in JPICLE (1, 2, 4). Given that, Turkish learners of English show wider use of *accept-n* patterns (6 patterns) than native speakers of English (5 patterns). This might be a

counterpart finding that Alqarni (2019) found that Swedish learners of English showed narrower use of patterns of *allow*. Unlike the findings of Altenberg and Granger (2001) that revealed learners of English from different language backgrounds abstained from using high-frequency verbs, Turkish learners of English showed a wider use in their essays. The patterns frequencies of learner corpora showed that learners with different proficiency levels and language backgrounds used fewer verb-noun patterns than native speakers of English (Laufer & Waldman, 2011). Further, it was seen that learners did not use pattern 5 and 6 of *allow* while native speakers also used these patterns. The possible reason of this finding may stem from the fact that Turkish and Japanese are postposition languages that differs from English which is a preposition language. Similar to the findings of Bağcı (2014), the verb-noun pattern use of Turkish learners showed their lexical collocational knowledge is better than their grammatical collocational knowledge (such as pattern 5 and 6 of *allow*).

4.5.2. Discussion of the Second Research Question

The investigation of the 7 verbs in learner corpora TICLE, NICLE, JPICLE and native corpus LOCNESS showed that learners produced some divergent patterns when the PDEV was taken as a reference for verb-noun collocation use of these learners. These possible divergent patterns were checked in the book *Grammar Patterns 1: Verbs* and searched in BNC to see their use in native language. The investigation showed that Turkish learners produced more possible divergent patterns ($f=9$) than Norwegian ($f=5$) and Japanese ($f=8$) learners of English. The variation of these numbers of patterns may show that the L1 of the foreign language learners affects their collocation use (Nesselnauf, 2003; Laufer & Waldman, 2011; Peters, 2016). Barfield (2006) found that Japanese learners had better knowledge in single verbs rather than the combination of verbs and nouns. Altenberg and Granger (2001) also highlighted the learner errors which are mostly occurred due to their L1. The divergent pattern use of Turkish and Japanese learners can be explained with the Yamashita and Jiang (2010) study which demonstrated that learners who had lower proficiency level in English tended to make more errors. However, even advance learners may not use verb-noun patterns correctly. There are some studies which supported the misuse or less use of collocations by advanced learners such as teachers (McCarthy, 1990; Zimmerman, 1997; Granger, 1998; Nesselnauf, 2003). L2 learners either over/underuse these patterns or misuse them which might result in collocational errors or divergent usages (Altenberg & Granger, 2001; Hasselgren, 1994; Nesselnauf, 2005; Bağcı, 2014). Nesselnauf (2005) indicated that these errors/divergent usages of patterns may stem from either the lack of pattern information in dictionaries or the learners' unawareness of the importance of collocations. However, the PDEV provides a reliable resource for the explanations of verb patterns and learners and teachers can easily make use of it to check collocations use. Furthermore, Franklin (2020) stated that the patterns and meaning of a verb were more important than the verb itself as the patterns of a specific verb may be used interchangeably. This supports the findings of this thesis in that learners do not

prefer some patterns of a verb as much frequent as other patterns of that verb that make some divergent pattern use. Finally, the divergent instances in learner corpora may be beneficial for the future intensification of the PDEV entries and the development of EFL classrooms to enhance the collocation awareness in English.

4.5.3. Discussion of the Third Research Question

This thesis was conducted by using Corpus Pattern Analysis (CPA) method and took the Pattern Dictionary of English Verbs (PDEV) as a reference tool to investigate the verb-noun patterns in learner corpora. Substantially, the PDEV was found as a reliable source to examine these patterns in target corpora although some instances could not be matched with the patterns existing in the PDEV while they had their patterns in the book *Grammar Patterns 1: Verbs* and their use in BNC. Alqarni (2019: 244) explained the lack patterns in learner data because of “the task the learners were asked to complete rather than a lack of knowledge”.

The over/underuse of the patterns showed the tendency of the target learners towards the specific patterns in the target verbs. For example, Turkish learners preferred using the pattern 1 of the verb *accept* 36 times to the pattern 3, 5 or 9. Moreover, the pattern 4 of the verb *allow* was highly frequent in LOCNESS ($f= 31$) and TICLE ($f= 12$) while NICLE ($f= 5$) and JPICLE ($f= 6$) had less frequency. The possible reason for the representation of the target verb patterns may arise from either the verb pattern itself or the learner’s knowledge about the target verb. Alqarni (2019) indicated that the frequency of verb may be low in native corpus which makes it uncommon pattern in learner data. On the other hand, Rees (2018: 211) stated that CPA/PDEV is a beneficial way to examine general language. Unlike idiom principle (Sinclair, 1991) and pattern grammar (Hunston & Francis, 2000) which focus on the structural features of language, CPA analyses language patterns from large sets of texts to help researchers understand language patterns based on the real language data. CPA/PDEV can be used to identify information about how verbs are frequently used with specific constructions. This approach to study language emphasized that language learners either adhere to existed patterns to convey meaning or deviate from those patterns to express meaning unconventionally (Hanks, 2013). It is undoubtedly possible that learner corpora can be identified through CPA as its main purpose is to investigate typical patterns of verbs and “to link word use and word meaning in a machine tractable way” (Hanks & Pustejovsky, 2005: 64). The findings of this thesis indicate that CPA/PDEV can be used to identify learner corpora in general; however, it is still difficult to explain the reason why learners use some pattern frequently while other patterns are used less frequently or not at all.

4.5.4. Discussion of the Fourth Research Question

Semantic types were included in these patterns to investigate the distinction between patterns and their meanings (Franklin, 2020). The verb patterns in the PDEV consist of the node word as well as the additional elements such as collocations, colligations, semantic and grammatical information of the verb. For example, the 20 pattern entries of the verb *follow* (see Appendix 6) includes different elements such as semantic information (Eventuality) and structural information (prepositional phrase or that clause). On the other hand, implications helped researcher to identify the patterns according to their meanings given in these entries. The PDEV also help identifying the learner errors in verb pattern use by applying CPA method. Hanks and Ma (2020) explained the effectiveness of the PDEV as a dictionary of English verbs that it uses the authentic data to demonstrate learners how verbs are used in native language. CPA/PDEV includes both syntactic and semantic analysis based on corpus data. The former investigates the idiomatic expressions, passive voice, phrasal uses, and the latter investigates the semantic features shared by collocations (El Maarouf & Baisa, 2013). Previous literature found that CPA is large enough to sample the general language which is not based on a specific discourse (Rees, 2018: 211; Franklin, 2020). Franklin (2020) similarly applied CPA to make discourse analysis of animal killing verbs and found the effectiveness of CPA as a reliable source to analyse the concordance lines easily. Thus, it can be inferred from their perspective that PDEV might also be used to describe learner production of English verb patterns and compare their use with native data.

In this thesis, three learner corpora were used to investigate how learners of English from different language backgrounds use verb-noun patterns. LOCNESS was used as a reference corpus to compare the learner and native language. The analyses were done by matching instances with their patterns in the PDEV which was used as second reference tool. It was found that PDEV is useful tool to investigate these patterns and possible divergent patterns in learner corpora; however, it might be insufficient in some points as it is still-growing tool and does not include all English verbs.

CONCLUSION

The aim of this thesis was to investigate the use of the verb-noun patterns by EFL learners from different language backgrounds (Turkish, Norwegian, Japanese) by applying Corpus Pattern Analysis (CPA) method which provides a highly informative and beneficial asset for computational linguistics which builds upon the Sinclairian approach to lexical analysis by examining the English verb use getting evidence from the British National Corpus (BNC) (Hanks, 2011; Popescu, 2012). The thesis also included the features of contrastive interlanguage analysis method as it covered three different learner corpora and a native corpus. As the learner data, sub-corpora of ICLE (TICLE, NICLE, JPICLE) was compared with the native data (LOCNESS) and a reference pattern dictionary (PDEV). The focus of this thesis was to examine whether EFL learners from different L1 backgrounds differ in using verb-noun patterns and how they produced these patterns in their academic essays. With this aim, the statistical and qualitative analysis were done to better understand the processes towards learner production.

This thesis constitutes of five chapters. Initially, *Introduction* tied together the concepts which are used in this thesis. It also presented the background, significance and organisation of the thesis and introduces the research questions. Secondly, *Literature Review*, initially introduced the notion of pattern, and continued to other concepts related to CPA and CIA. The importance of learner corpus research and error analysis were discussed in this chapter, and previous literature were provided to show relevant studies to this thesis. The third chapter, *Methodology*, in which the research design of the thesis was discussed, explains the learner corpora (ICLE) and the native corpus (LOCNESS) used in the thesis; data analysis tools (Sketch Engine and the PDEV). This chapter was summarised providing information about the research steps to be taken in this thesis. The fourth chapter, *Findings and Discussion*, provided the statistical analysis and their detailed investigation. Lastly, this chapter is a general summary of the thesis and the findings. The summary will be followed by the limitations of the thesis, pedagogical implications, and further suggestions.

The selection of target verbs is based on the collocation list of Pearson Test of English (PTE) which is a set of word groups to help English learners using the right word combinations to be successful in exams. This list covers the collocations frequently occur in the PTE exam which comprise of common collocations around a key word or a base word. In this study, of the 2,469 most frequent collocations in English, 63 verb-noun collocations were found. However, mere 21 verbs were mutual in each corpus in this thesis. 7 verbs were chosen for further analysis as they had frequency in each source such as ICLE, LOCNESS and PDEV. The data was uploaded to Sketch

Engine to generate the concordance lines and to survey their frequency. These concordance lines were matched with their patterns in the PDEV. Later, their frequency tables and log-likelihood value tables demonstrated the similarities and differences of verb-noun pattern use between corpora. The statistical analysis showed that two verbs *accept* and *allow* were statistically significant to further investigate. The findings showed that Turkish learners of English showed a wider verb-noun pattern use than Norwegian and Japanese learners. On the one hand, it was observed in Turkish and Japanese corpora that they tend to produce more divergent patterns than Norwegian learners; on the other hand, the possible reason for the similarity between Turkish and Japanese learners and difference between these learners and Norwegian learners is their proficiency levels.

The findings reveals that some verbs are over or underused in learner corpora with a similar LL value which shows the frequency of collocations of the target verbs with certain words and their statistically significant difference. Only, the verb *attend* is found overused in each corpus when their frequency is compared to LOCNESS frequency. The differences between the use of verb-noun patterns between corpora indicated that each learner from different language backgrounds may produce particular sets of verb patterns less frequent as their knowledge for a same verb is likely to lack (Barfield, 2006). In addition, some patterns could not be found in the PDEV; thus, these instances were classified as possible divergent patterns which could be turned into error later. At this point, CPA and PDEV demonstrates a significant description of learner corpora. To analyse the pattern not found in the PDEV, corpus-based interlanguage analysis was implemented to understand the L2 production of language learners. These unknown patterns may be the indicative of L2 language proficiency and learner difficulties in language production (Ortega, 2009). Some verb-noun usages which are not found in BNC may also show that they might be L1-related if learners translate multi-word expressions directly from their native language into English (Paquot et al., 2006).

Limitations of the Thesis

CPA and PDEV are successful ways of interpreting learner language; however, PDEV is not complete and does not include patterns for all English verbs. Thus, this study is limited to certain verbs as others do not have their patterns in the PDEV. Additionally, working on three learner corpora is another limitation of this thesis as it is difficult to find frequency of each verb in these corpora.

Implications

The findings of this thesis provide several implications including pedagogical implications. It is important to bear in mind that learner corpora provide valuable insights to language researcher in

discerning the trajectory of lexical acquisition of language learners (Özbay, 2015). Unlike various learner corpus studies about verb-noun collocations, this thesis is important as being the only study in Türkiye using the newly developed tool PDEV to investigate the learner corpora.

Pedagogical Implications

It is important to study learner language with a contrastive methodology to understand the underlying systems of the rules of learner language and its reflections to L2 production. Teaching grammatical patterns to language learners such as verb-noun patterns might increase both learner L2 awareness and their proficiency in English. As a result of the analysis of three learner data (TICLE, NICLE, JPICLE), this thesis provides some pedagogical implications for English classrooms by proposing CPA/PDEV.

First of all, CPA is a corpus-based method to analyse pattern, and its outcome PDEV is based on corpus data. Thus, both might be related to Data-Driven Learning (DDL) activities which is a term coined by Johns (1990) as a popular teaching and learning method in English Language Teaching which helps learners explore the patterns and other features of language themselves. Boulton (2012) stated that there are online corpus-based tools which might provide instant benefits to language learners through DDL activities. The application of DDL requires teacher-created concordances based on corpus data to provide learners to comprehend language structures such as vocabulary, grammar and collocations (Croswaithe et al., 2021: 1). Language teacher may use DDL to teach verb-noun patterns to their learners with various classroom activities (Pinto et al., 2023). Thus, learners can easily discover language patterns based on authentic language data with or without the help of their teachers (Boulton, 2007: 1).

The application of CPA method enhances the communicative skills of learners, strengthen their grammatical and vocabulary knowledge, and encourage them to use English by adopting the learner-centered methods. Through CPA, learners can observe the rules and the concepts of the language and learn how to combine their knowledge with real-life applications. Moreover, the analysis of real language patterns enables students to understand how language can change in different contexts.

As an outcome of CPA method, Pattern Dictionary of English Verbs (PDEV) is a lexicon-based tool which shows the frequently occurred verb patterns based on the BNC50 data. It specifically focuses on the verb patterns of specific verbs and their grammatical structures. Students can learn the correct use of verbs in their written and oral communication by observing these patterns in the PDEV. The tool can also be used to detect learner errors in verb use through learner corpus analysis. These errors might help language teachers to understand whether or how learners have difficulty in understanding and using verbs in their learning process. Thus, teachers

can apply new methods or strategies to overcome these problematic language use. All in all, both methods help learners to understand grammatical patterns of English and language teachers to develop new materials to teach these patterns.

Further Suggestions

This thesis focused on verb-noun patterns in Turkish, Norwegian and Japanese written corpora. As it is detailed in the findings, only 7 verbs were chosen for analysis and 2 verbs are chosen for further analysis. It is recommended that further research to be undertaken in this field may choose more language backgrounds as alternative to compare the data to see what the similarities and differences between other languages are. Although the PDEV is not complete and growing source for CPA method, it can be used in EFL classrooms to encourage learners to understand how some certain verbs come and gain meaning together. CPA enables analysing language patterns using real language data which helps learners to observe authentic language use and relate it to the real world. As students see real language examples, they gain a better understanding of how language is used in real life and how it can change in different contexts. Another important issue which helps future studies that PDEV can contribute to the development of teaching resources and materials based on the results of learner corpus analysis. Teaching materials can be provided to help learners understand and use verb patterns.

REFERENCES

- Akıncı, Mehmet (2009), **Effectiveness of Corpus Consultancy in Teaching English Verb Noun Collocations to First Year ELT Students**, Unpublished Master's Thesis, Boğaziçi University.
- Allen, David (2011), "Lexical Bundles in Learner Writing: An Analysis of Formulaic Language in the ALESS Learner Corpus", **Komaba Journal of English Education**, 105-127.
- Alqarni, Abdullah Saad (2019), **Applying Corpus Pattern Analysis to Learner Corpora: Investigating the Pedagogical Potential of the Pattern Dictionary of English verbs**, Unpublished PhD Thesis, University of Birmingham.
- Altenberg, Bengt and Granger, Sylviane (2001), "The Grammatical and Lexical Patterning of MAKE in Native and Non-Native Student Writing", **Applied Linguistics**, 22(2), 173–195. <https://doi.org/10.1093/applin/22.2.173>
- Babanoğlu, M. Pınar (2012), **A Corpus-Based Study on Turkish EFL Learners' Written English: The Use of Adverbial Connectors by Turkish Learners**. Unpublished PhD Thesis, Çukurova University
- Bağcı, Nazife Duygu (2014), **Turkish University Level EFL Learners' Collocational Knowledge at Receptive and Productive Levels**. Unpublished Master's Thesis, Middle East Technical University.
- Barfield, Andrew William (2006), **An Exploration of Second Language Collocation Knowledge and Development**. Unpublished PhD Thesis, University of Wales Swansea.
- Barnbrook, Geoff (2007), "Sinclair on Collocation", **International Journal of Corpus Linguistics**, 12(2), 183-199. <https://doi.org/10.1075/ijcl.12.2.06bar>
- Boers, Frank et al. (2006), "Formulaic Sequences and Perceived Oral Proficiency: Putting a Lexical Approach to the Test", **Language Teaching Research**, 10(3), 245-261. <https://doi.org/10.1191/1362168806lr195oa>
- Boulton, A. (2007). "DDL is in the details...and in the Big Themes", In M. Davies, P. Rayson, S. Hunston, & P. Danielsson (Eds.), **Proceedings of Corpus Linguistics**.
- Boulton, Alex. (2012), "What Data for Data-Driven Learning?", **EuroCALL Review**, 23-27. <http://dx.doi.org/10.4995/eurocall.2012.16038>

- Bradbury, Jane and El Maarouf, Ismail (2013, November), An empirical classification of verbs based on Semantic Types: the case of the 'poison' verbs. In Proceedings of the joint symposium on semantic processing. Textual inference and structures in corpora (pp. 70-74).
- Brown, Henry Douglas (2007), **Principles of Language Learning and Teaching**, 5th Ed., Longman, New York.
- Can, Cem (2017), "A Learner Corpus-Based Study on Verb Errors of Turkish EFL Learners", **Journal of Education and Training Studies**, 5(9), 167-175. <https://doi.org/10.11114/jets.v5i9.2612>
- _____ (2018), "Agreement Errors in Learner Corpora Across CEFR: A Computer-Aided Error Analysis of Greek and Turkish EFL Learners Written Productions", **Journal of Education and Training Studies**, 6(5), 77-84. <http://dx.doi.org/10.11114/jets.v6i5.3064>
- Cinková, Silvie and Hanks, Patrick (2010), "Validation of Corpus Pattern Analysis - Assigning pattern numbers to random verb samples" Retrieved from http://nlp.fi.muni.cz/projects/cpa/CPA_valiman.pdf
- Corder, Stephen Pit (1967), "The Significance of Learner's Errors", **IRAL: International Review of Applied Linguistics in Language Teaching**, 5(1-4), 161-170. <https://doi.org/10.1515/iral.1967.5.1-4.161>
- _____ (1982), **Error Analysis and Interlanguage**, 2nd Ed., Oxford University Press.
- Creswell, John W. (2009), **Research Design: Qualitative, Quantitative and Mixed Method Approaches**, 3rd Ed., SAGE Publications.
- Crosthwaite, Peter et al. (2021), "Exploring language teachers' lesson planning for corpus-based language teaching: a focus on developing TPACK for corpora and DDL", **Computer Assisted Language Learning**, 1-29. <https://doi.org/10.1080/09588221.2021.1995001>
- Edwards, Alison and Laporte, Samantha (2015), "Outer and Expanding Circle Englishes: The Competing Roles of Norm Orientation and Proficiency Levels", **English World-Wide**, 36(2), 135-169. <https://doi.org/10.1075/eww.36.2.01edw>
- Edwards, Alison and Lange, Ruther Jan (2016), "In Case of Innovation", **International Journal of Learner Corpus Research**, 2(2), 252-277. <https://doi.org/10.1075/ijlcr.2.2.06edw>
- Ellis, Rod (1994), **The Study of Second Language Acquisition**, Oxford University Press.
- Ellis, Nick C. et al., (2015), "Learner Corpora and Formulaic Language in Second Language Acquisition", S. Granger, G. Gilquin, & F. Meunier (Eds.), In **the Cambridge Handbook of Learner Corpus Research** (357-378). Cambridge University Press, Cambridge. <https://10.1017/CBO9781139649414.016>

- El Maarouf, Ismail (2013), "Methodological Aspects of Corpus Pattern Analysis", **ICAME**, 37, 119- 148.
- El Maarouf, Ismail and Baisa, Vít (2013), "Automatic classification of semantic patterns from the Pattern Dictionary of English Verbs". In **Proceedings of the Joint Symposium on Semantic Processing. Textual Inference and Structures in Corpora**, 95-99, Trento, Italy.
- El Maarouf, Ismail et al. (2014a), "Disambiguating Verbs by Collocation: Corpus Lexicography meets Natural Language Processing". In **LREC**, 1001- 1006.
- El Maarouf, Ismail et al. (2014b), "PDEV-lemon: a Linked Data implementation of the Pattern Dictionary of English Verbs based on the Lemon model". In **3rd Workshop on Linked Data in Linguistics: Multilingual Knowledge Resources and Natural Language Processing**, 88-93.
- Erman, Britt et al., (2016), "Formulaic Language in Advanced Second Language Acquisition and Use", Kenneth Hyldenstam (Ed.), In **Advanced Proficiency and Exceptional Ability in Second Languages**, (111-147), De Gruyter Mouton. <https://doi.org/10.1515/9781614515173-007>
- Firth, John Rupert (1957), **Papers in Linguistics**, Oxford University Press, Oxford.
- Flowerdew, Lynne (2015), "Learner Corpora and Language for Academic and Specific Purposes", Sylviane Granger, Gaëtanelle Gilquin, & Fanny Meunier (Eds.), In **The Cambridge Handbook of Learner Corpus Research**, (465–484), Cambridge University Press, Cambridge. <http://dx.doi.org/10.1017/CBO9781139649414.021>
- Francis, Gill (1993), "A Corpus-Driven Approach to Grammar: Principles, Methods and Examples", Mona Baker, Gill Francis, Elena Tognini-Bonelli (Eds), In **Text And Technology: In Honour Of John Sinclair**, (137-156), John Benjamins Publishing Company. <https://doi.org/10.1075/z.64>
- Francis, Gill et al. (1996), **Collins Cobuild Grammar Patterns 1: Verbs**, Harper Collins, London.
- Franklin, Emma (2020), **Acts of Killing, Acts of Meaning: An Application of Corpus Pattern Analysis to Language of Animal-Killing**, Unpublished PhD Thesis, Lancaster University.
- Gablasova, Dana et al. (2017), "Collocations in Corpus-Based Language Learning Research: Identifying, Comparing, and Interpreting the Evidence", **Language Learning**, (155-179). <https://doi.org/10.1111/lang.12225>
- Gençer, Bora (2004), **Raising EFL Learners' Awareness of Verb-Noun Collocations Through Chunking to Extend Their Collocational Knowledge of Familiar Nouns**, Unpublished Master's Thesis, Anadolu Üniversitesi – Eğitim Bilimleri Enstitüsü.

- Gilquin, Gaëtanelle et al. (2007), "Learner Corpora: The Missing Link in EAP Pedagogy", **Journal of English for Academic Purposes**, 6(4), 319-335.
<https://doi.org/10.1016/j.jeap.2007.09.007>
- Granger, Sylviane (1996), "From CA To CIA And Back: An Integrated Contrastive Approach to Computerized Bilingual And Learner Corpora", Karin Aijmer, Bengt Altenberg, Mats Johansson (Ed), In **Languages In Contrast. Text-Based Cross-Linguistic Studies**, (37-51), Lund University Press, Lund.
- _____ (1998), "The Computer Learner Corpus: A Versatile New Source of Data For SLA Research" Sylviane Granger (Ed), In **Learner English on Computer**, (3-18), Addison Wesley Longman, London & New York.
- _____ (2015), "Contrastive Interlanguage Analysis: A Reappraisal", **International Journal of Learner Corpus Research**, 1(1), 7-24. <https://doi.org/10.1075/ijlcr.1.1.01gra>
- _____ (2019), "Learner Corpora", Carol Ann Chapelle (Ed), In **The Encyclopaedia of Applied Linguistics**, (681-688), Wiley Blackwell.
- Granger, Sylviane and Tyson, Stephanie (1996), "Connector Usage in The English Essay Writing of Native and Non- Native EFL Speakers of English" **World Englishes**, 15(1), 17-27.
<https://doi.org/10.1111/j.1467-971X.1996.tb00089.x>
- Granger, Sylviane et al. (2006), "Extraction of Multi-Word Units from EFL and Native English Corpora. The Phraseology of the Verb 'Make'", In A. Häcki Buhofer & H. Burger (Eds.), *Phraseology in motion I: Methoden und Kritik* (57–68). Akten der Internationalen Tagung zur Phraseologie. Baltmannsweiler, Germany: Schneider Verlag Hohengehren.
- Granger, Sylviane et al. (2020), **The International Corpus of Learner English**, 3rd Ed., Presses Universitaires de Louvain, Louvain-la-Neuve.
- Hanks, Patrick and Pustejovsky, James (2005), "A Pattern Dictionary for Natural Language Processing", **Revue Française De Linguistique Appliquée**, 10(2), 63-82.
<https://doi.org/10.3917/rfla.102.82>
- Hanks, Patrick and Bradbury, Jane (2013), "Why Do We Need Pattern Dictionaries (And What Is a Pattern Dictionary, Anyway)", **Kernerman Dictionary News**, 21, 27-31.
- Hanks, Patrick and Može, Sara (2019), "The Way to Analyse "Way": A Case Study in Word-Specific Local Grammar", **International Journal of Lexicography**, 32(3), 247-269.
<https://doi.org/10.1093/ijl/ecz005>
- Hanks, Patrick and Ma, Wulin (2020), "Meaning and Grammar in The Light of Corpus Pattern Analysis", **International Journal of Lexicography**, 34(1), 135-149.
<https://doi.org/10.1093/ijl/ecaa027>

- Hanks, Patrick (2004), "Corpus Pattern Analysis", Geoffrey Williams, Sandra Vessier (Eds.), **11th Euralex International Congress Proceedings**, Lorient, Université de Bretagne-Sud, 87-97.
- _____ (2012), "Corpus Evidence And Electronic Lexicography", Sylviane Granger, Magali Paquot (Eds.), In **Electronic Lexicography**, (57-82), Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199654864.001.0001>
- _____ (2013), **Lexical Analysis: Norms and Exploitations**, MIT Press. <https://doi.org/10.7551/mitpress/9780262018579.001.0001>
- Hasselgård, Hilde and Johansson, Stig (2011), "Learner Corpora and Contrastive Interlanguage Analysis", F. Meunier, S. De Cock, G. Gilquin & M. Paquot (Eds.), In **A Taste for Corpora in Honour of Sylviane Granger**, (33-61) Amsterdam and Philadelphia: John Benjamins. <https://doi.org/10.1075/scl.45.06has>
- Hasselgren, A. (1994), "Lexical Teddy Bears and Advanced Learners: A Study into The Ways Norwegian Students Cope with English Vocabulary", **International Journal of Applied Linguistics**, 4(2), 237-258.
- Hoey, Michael (2005), **Lexical Priming: A New Theory of Words and Language**, Routledge, London.
- Hornby, Albert Sydney (1954), **A Guide to Patterns and Usage in English**, Oxford University Press.
- _____ (1975), **A Guide to Patterns and Usage in English**, 2nd Ed., Oxford University Press.
- Hill, Jimmie (1999), "Collocational Competence", **Readings in Methodology**, 162-166.
- Hunston, Susan (2015), "Lexical Grammar", Douglas Biber and Randi Reppen (Eds.), In **The Cambridge Handbook of English Corpus Linguistics**, (201-215). Cambridge University Press. <https://doi.org/10.1017/CBO9781139764377>
- Hunston, Susan (2022), **Corpora in Applied Linguistics**, 2nd Ed., Cambridge University Press. <https://doi.org/10.1017/9781108616218>
- Hunston, Susan and Francis, Gill (2000), **Pattern Grammar: A Corpus-Driven Approach to The Lexical Grammar of English**, John Benjamins Publishing. <https://doi.org/10.1075/scl.4>
- Hunston, Susan and Su, Hang (2017), "Pattern, Constructions, And Local Grammar: A Case Study of 'Evaluation'", **Applied Linguistics**, 40(4), 567-593. <https://doi.org/10.1093/applin/amx046>

- Jeckins, Jennifer et al. (2011), "Review of Developments in Research into English as A Lingua Franca", **Language Teaching**, 44(3), 281-315.
<https://doi.org/10.1017/S0261444811000115>
- Johns, Tim. (1990), "From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning", **CALL Austria**, 10, 14-34.
- Kırkgöz, Yasemin (2010), "An analysis of written errors of Turkish adult learners of English" **Procedia- Social and Behavioral Sciences**, 2(2), 4352-4358.
<https://doi.org/10.1016/j.sbspro.2010.03.692>
- Kilgariff, Adam et al. (2004), "The Sketch Engine", **Proceedings of the 11th EURALEX International Congress**, 105-116.
- Lakshmanan, Usha and Selinker, Larry (2001), "Analysing Interlanguage: How do we know what learners know?", **Second Language Research**, 17, 393-420.
- Laufer, Batia and Waldman, Tina (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>
- Lewis, Michael (2000), **Teaching Collocations: Further Developments in The Lexical Approach**, Language Teaching Publications, Hove.
- Littlemore, Jeannette (2015), **Metonymy: Hidden Shortcuts in Language, Thought and Communication**, Cambridge University Press, Cambridge.
- López, Willelmira Castillejos (2009), "Error Analysis in A Learner Corpus: What Are the Learners' Strategies", **A Survey of Corpus-Based Research**, 675-690.
- McCarthy, Michael (1990), **Vocabulary**, Oxford University Press, Oxford.
- McCarthy, Michael and O'Dell, Felicity (2005), **English Collocations in Use**, Cambridge University Press, Cambridge.
- McCarthy, Michael and O'Dell, Felicity (2017), **English Collocations in Use**, 2nd Ed., Cambridge University Press, Cambridge.
- McEnery, Tony et al. (2006), **Corpus-based Language Studies: An Advanced Resource Book**, Routledge Applied Linguistics, London.
- McEnery, Tony and Brezina Vaclav (2022), **Fundamental Principles of Corpus Linguistics**, Cambridge University Press, Cambridge. <https://doi.org/10.1017/9781107110625>
- Nesselhauf, Nadja (2003), "The Use of Collocations By Advanced Learners of English And Some Implications For Teaching", **Applied Linguistics**, 24(2), 223-242.
<https://doi.org/10.1093/applin/24.2.223>

- _____ (2005), **Collocations in A Learner Corpus**, John Benjamins.
<https://doi.org/10.1075/scl.14>
- Ortega, Lourdes. (2009), **Understanding Second Language Acquisition**. Routledge.
- Oxford Collocations Dictionary for Students of English. (2002). *Oxford University Press*.
- Özbay, Ali Şükrü. (2015), **Corpus Analysis of the Support Verb Construction Development and Use by EFL Learners**, Unpublished PhD Dissertation, Karadeniz Technical University.
- Peters, Elke (2016) “The learning burden of collocations: The role of interlexical and intralexical factors”, **Language Teaching Research**, 20(1), 113-138.
<https://doi.org/10.1177/1362168814568131>
- Pinto, Paula Tavares et al. (2023), **Using Language Data to Learn About Language: A Teachers' Guide to Classroom Corpus Use**. Brisbane, QLD Australia: The University of Queensland. <https://doi.org/10.14264/3bbe92d>
- Popescu, Octavian. (2012), “Building a Resource of Patterns Using Semantic Types”, **International Conference on Language Resources and Evaluation**, 2999-3006.
- Popescu, Octavian et al. (2014), “Mapping CPA Patterns onto OntoNotes Senses”, **International Conference on Language Resources and Evaluation**, 882-889.
- Rees, Geraint Paul (2018), **A Phraseological Multi-Discipline Approach to Vocabulary Selection for English for Academic Purposes**, Unpublished PhD Thesis, Universitat Pompeu Fabra.
- Rees, Geraint Paul (2021), “Discipline-Specific Academic Phraseology”, Maggie Charles, Ana Frankenberg-Garcia (Eds.), In **Corpora in ESP/EAP Writing Instruction**, (32-54). Routledge. <http://dx.doi.org/10.4324/9781003001966-2-4>
- Richards, Jack Croft and Schmidt, Richard (2002), **Dictionary of Language Teaching & Applied Linguistics: Pearson Education Limited**, Longman, London.
- Salazar, Danica (2014), **Lexical Bundles in Native and Non-native Scientific Writing**, John Benjamins, Amsterdam. <https://doi.org/10.1075/scl.65>
- Schneider, Gerold and Gilquin, Gaëtanelle (2016), “Detecting Innovations in A Parsed Corpus Of Learner English”, **International Journal of Learner Corpus Research**, 2(2), 177-204.
<https://doi.org/10.1075/ijlcr.2.2.03sch>
- Seidlhofer, Barbara (2007), “Common Property: English as a Lingua Franca in Europe”, Jim Cummins, Chris Davison (Eds.), In **International Handbook of English Language Teaching**, (137-153). Springer, New York. http://dx.doi.org/10.1007/978-0-387-46301-8_11

- Selinker, Larry (1972), "Interlanguage", **International Review of Applied Linguistics in Language Teaching**, 10(3).
- Sinclair, John (1991), **Corpus, Concordance, Collocation**, Oxford University Press, Oxford.
- _____ (2004), **Trust The Text: Language, Corpus, And Discourse**, Routledge.
- Tarone, Elaine (2006), "Interlanguage", Keith Brown (Ed.), In **Encyclopedia of Language and Linguistics**, (747-751), Elsevier, Boston.
- Teddlie, Charles and Tashakkori, Abbas (2008), **Foundations of Mixed Methods Research**. Thousand Oaks CA: Sage.
- URL, <http://www.sketchengine.eu/>, Access Date, 16.02.2023.
- Yamashita, Junko and Jiang, Nan. (2010), "L1 Influence on the Acquisition of L2 Collocations: Japanese ESL Users and EFL Learners Acquiring English Collocations", **TESOL Quarterly**, 44(4), 647-668. <https://doi.org/10.5054/tq.2010.235998>
- Yıldız, Mustafa and Turan, Ümit Deniz (2021), "Contrastive Interlanguage Analysis of Evidentiality in PhD Dissertations", **Discourse and Interaction**, 14(1), 124-152. <https://doi.org/10.5817/DI2021-1-124>
- Zimmerman, Cheryl Boyd (1997), "Historical Trends in Second Language Vocabulary Instruction" James Coady, Thomas. Huckin (Eds.), In **Second Language Vocabulary Acquisition** (5-19). Cambridge: Cambridge University Press.



APPENDICES

Appendix 1: Verb Pattern Entries of *accept* in PDEV

Pattern Number	Pattern and Implicature
1	Pattern: [[Human Institution]] accept [[Proposition Concept Eventuality]]
	Implicature: [[Human Institution]] agrees that [[Proposition Concept Eventuality]] is correct and does not need to be contested
2	Pattern: [[Human Institution]] accept
	Implicature: [[Human Institution]] agrees that {that [CLAUSE]} is true or correct
3	Pattern: [[Human Institution]] accept {responsibility} [for [[Eventuality Entity]]]
	Implicature: [[Human Institution]] agrees that he, she, or it is or will be responsible for [[Eventuality Entity]]
4	Pattern: [[Human 1]] accept [[Entity = Valuable Money]]
	Implicature: [[Human 1]] consents to receive [{Entity = Valuable} Money]] as a gift (or bribe) from [[Human 2]]
5	Pattern: [[Human]] accept [[Offer]] {offer proposal recommendation suggestion advice}
	Implicature: [[Human 1]] agrees to act on the [[Offer]] of [[Human 2]]
6	Pattern: [[Human 1 = Authority]] accept {resignation}
	Implicature: [[Human 1 = Authority]] receives and consents to {resignation} of [[Human 2 = Employee]]
7	Pattern: [[Human]] accept {appointment job}
	Implicature: [[Human]] agrees to undertake the work specified as required to fulfil {appointment job}
8	Pattern: [[Human 1 Institution 1]] accept [[Human 2 Institution 2]] [as [[Human_Role Institution_Role]]]
	Implicature: [[Human 1 Institution 1]] acknowledges that [[Human 2 Institution 2]] is suitable for [[Human_Role Institution_Role]]
9	Pattern: [[Human_Group 1 Human 1]] accept [[Human 2 Human_Group 2]]
	Implicature: [[Human_Group 1 Human 1]] agrees to allow [[Human 2 Human_Group 2]] to become a member of or be respected by [[Human_Group 1]]

Appendix 2: Verb Pattern Entries of *add* in PDEV

Pattern Number	Pattern and Implicature
1	Pattern: [[Human] add [QUOTE] [THAT]
	Implicature: [[Human]] says or writes something that gives additional information
2	Pattern: [[Human]] add [[Physical _ Object 1]] [to [[Physical _ Object 2]]]
	Implicature: [[Human]] puts [[Physical _ Object 1]] with, alongside, or into [[Physical Object 2]]
3	Pattern: [[Human Institution Eventuality 1]] add [Abstract _ Entity 1 Eventuality 2]] [on,to [[State_of_Affairs Abstract_Entity 2]]]
	Implicature: [[Human Institution Eventuality 1]] contributes [[Abstract _ Entity 1 Eventuality 2]] ({to [[State _of Affairs]] to [Abstract Entity 2]})
4	Pattern: [[Eventuality]] add [to [[State_of_Affairs = Existing]]]
	Implicature: [[Eventuality]] contributes to [[State_of_Affairs = Existing]]
5	Pattern: [Human Eventuality]] add [[Numerical _ Value 1]] [to [Numerical Value 2]]]
	Implicature: [[Human Eventuality]] causes [[Numerical _ Value 2]] to increase by [[Numerical Value 1]]
6	Pattern: phrasal [[Human]] add [[Numerical _ Value 1 = Plural]] [up, together]
	Implicature: [[Human]] combines [[Numerical _ Value 1 = Plural]] to produce [[Numerical Value 2]]
7	Pattern: phrasal [[Numerical _ Value 1 = Plural]] add [up, to [[Numerical _ Value 2]]]
	Implicature: [[Numerical _ Value 1 = Plural]] is combined to produce [[Numerical Value 2]]}
8	Pattern: phrasal [[Eventuality 1]] add [up, to [[Eventuality 2]]]
	Implicature: All the factors involved in [[Eventuality 1]] are considered to be the equivalent of [[Eventuality 2]]

Appendix 3: Verb Pattern Entries of *affect* in PDEV

Pattern Number	Pattern and Impicature
1	Pattern: [[Eventuality 1 Entity 1]] affect [[Human Animal Eventuality 2 Entity 2]]
	Impicature: [[Eventuality 1 Entity 1]] causes [[Human Animal Eventuality 2 Entity 2]] to change
2	Pattern: [[Human]] affect [[Attitude Emotion]]
	Impicature: [[Human]] pretends to have [[Attitude]] or to feel [[Emotion]]

Appendix 4: Verb Pattern Entries of *allow* in PDEV

Pattern Number	Pattern and Implicature
1	Pattern: [[Human 1 Institution 1 Eventuality]] allow [[Human 2 Institution 2 Self]] [to+INF]
	Implicature: [[Human 1 Institution 1 Eventuality]] gives [[Human 2 Institution 2 Self]] the opportunity, time, or permission {to/INF [V]}
2	Pattern: [[Human 1 Institution 1 Eventuality]] allow [[Institution 2 Self Privilege Human 2]]
	Implicature: [[Human 1 Institution 1 Eventuality]] gives [[Human 2 Institution 2 Self]] permission or the opportunity to have [[Privilege]]
3	Pattern: [[Eventuality Human Institution]] allow [[Physical_Object]] {to be V-ED} [to+INF]
	Implicature: [[Human Institution Eventuality]] causes or provides the opportunity for [[Physical_Object]] {to be V-ED}
4	Pattern: [[Human Eventuality 1 Institution]] allow [[Eventuality 2]]
	Implicature: [[Human Institution Eventuality 1]] gives the opportunity, time, or permission for [[Eventuality 2]] to occur
5	Pattern: [[Human 1 Institution 1]] allow [[Human 2 Animal]] [Adv [Direction]]
	Implicature: [[Human 1 Institution 1]] gives permission for [[Human 2 Animal]] to move in [Adv [Direction]]
6	Pattern: [[Human]] allow [[Resource]] [for [[Activity]]]
	Implicature: When planning, [[Human]] ensures that a sufficient amount of [[Resource]] is made available to complete [[Activity]]
7	Pattern: [[Human Eventuality 1]] allow [for [[Eventuality 2]]]
	Implicature: [[Human Eventuality 1]] makes provision {for [[Eventuality 2]]}
8	Pattern: [[Human]] allow [THAT]
	Implicature: [[Human]] admits the truth of {that-CLAUSE}

Appendix 5: Verb Pattern Entries of *attend* in PDEV

Pattern Number	Pattern and Implicature
1	Pattern: [[Human]] attend ([[Event]]) [[Location = <i>Functional</i>]]) {school college court}
	Implicature: [[Human]] is present at [[Event Location = Functional]]
2	Pattern: [Human]] attend [to [[Anything]]]
	Implicature: [[Human]] pays attention to and deals as necessary with [[Anything]]
3	Pattern: [[Human 1]] attend [[Human 2]]
	Implicature: [[Human 1]] looks after [[Human 2]]



Appendix 6: Verb Pattern Entries of *follow* in PDEV

Pattern Number	Pattern and Implicature
1	Pattern: [[Human 1 Animal 1 Vehicle 1]] follow [[Human 2 Animal 2 Vehicle 2]]
	Implicature: [[Human 1 Animal 1 Vehicle 1]] moves in the same direction as that selected by [[Human 2 Animal 2]] or the driver of [[Vehicle 2]]
2	Pattern: [[Human 1 = <i>Student</i> <i>Disciple</i>]] follow [[Human 2 = <i>Teacher</i>]]
	Implicature: [[Human 1 = Student Disciple]] studies and is influenced by or tries to practice the teachings of [[Human 2 = Teacher]]
3	Pattern: [[Human_Group 1 Institution 1]] follow [[Human_Group 2 Institution 2]]
	Implicature: [[Human_Group 1 Institution 1]] is influenced by and thus copies the behaviour of [[Human_Group 2 Institution 2]]
4	Pattern: [[Human 1]] follow [[Human 2]] [[Human_Role]]
	Implicature: [[Human 1]] takes on [[Human_Role]] after [[Human 2]] relinquishes it
5	Pattern: [[Institution Human]] follow [[Command Rule Plan Document]]
	Implicature: [[Human Institution]] acts in accordance with [[Command Rule Plan]] (expressed in [[Document]])
6	Pattern: [[Institution Human]] follow [[Event = <i>Unfolding</i>]]
	Implicature: [[Human Institution]] takes an interest in [[Event = Unfolding]]
7	Pattern: [[Human]] follow [[Proposition]] [WH+]
	Implicature: [[Human]] is able to understand {[[Proposition]] {WH-CLAUSE}}
8	Pattern: [[Human]] follow {trade}
	Implicature: [[Human]] earns a living by practising {[MOD] trade}
9	Pattern: [[Human Institution]] follow {lead line of enquiry}
	Implicature: [[Human Institution]] pursues an investigation based on specific information
10	Pattern: [[Human Institution]] follow {trend}
	Implicature: [[Human Institution]] performs in line with the expectations dictated by {trend}
11	Pattern: [[Event 1]] follow [[Event 2]]
	Implicature: [[Event 1]] happens after and typically as a consequence of [[Event 2]]
12	Pattern: [[Eventuality 1]] follow [from [[Eventuality 2]]]
	Implicature: [[Eventuality 1]] is a necessary consequence of [[Eventuality 2]]
13	Pattern: {it} follow [from [[Eventuality 1]] [THAT] [THAT]]
	Implicature: {[[Eventuality 2]] expressed by [CLAUSE]} is a necessary consequence of [[Eventuality 1]] (either expressed in a from-phrase or one that was previously stated)
14	Pattern: [Anything 1 = <i>Item in List</i>] follow [[Anything 2 = <i>Item in List</i>]]
	Implicature: [[Anything 1 = Item in List]] comes after [[Anything 2 = Item in List]] in a list or sequence
15	Pattern: phrasal [[Human]] follow [[Eventuality]] [up]
	Implicature: [[Human]] monitors [[Eventuality]] and finds out more about it
16	Pattern: phrasal [[Human]] follow [on, up [[Eventuality]]]
	Implicature: [[Human]] monitors [[Eventuality]] and finds out more about it
17	Pattern: idiom as follows
	Implicature: as will now be stated here
18	Pattern: idiom {[[Human 1 Institution 1]] follow {suit}}
	Implicature: [[Human 1 Institution 1]] does the same as [[Human 2 Institution 2]] has done
19	Pattern: idiom [[Human 1]] follow [in {footsteps (of [[Human 2]])}]
	Implicature: [[Human 1]] engages in similar activity to one previously engaged in by [[Human 2]] }
20	Pattern: idiom [[Human Institution]] be hard act to follow
	Implicature: it is difficult to emulate or surpass [[Human Institution]] because he, she, or it has performed so well

Appendix 7: Verb Pattern Entries of *maintain* in PDEV

Pattern Number	Pattern and Implicature
1	Pattern: [[Human Process]] maintain [[State_of_Affairs]]
	Implicature: [[Human]] takes action to ensure or [[Process]] has the effect that [[State_of_Affairs]] remains unchanged
2	Pattern: [[Human Institution Document]] maintain
	Implicature: [[Human Institution Document]] asserts strongly and consistently over time {that [CLAUSE] [Proposition]}
3	Pattern: [[Human Institution]] maintain [[Document]]
	Implicature: [[Human Institution]] regularly and systematically updates [[Document]]
4	Pattern: [[Human]] maintain [[Building Vehicle Artifact]]
	Implicature: [[Human]] takes action to ensure that [[Building Vehicle Artifact]] is kept in good working order
5	Pattern: [[Human 1]] maintain [[Self Human 2]]
	Implicature: [[Human 1]] provides food and drink for [[Self Human 2]], enabling [[Self Human 2]] to stay alive

CURRICULUM VITAE

She graduated from Çorlu Anatolian High School in 2015. Then, she graduated from the Department of English Language and Literature at Karadeniz Technical University in 2020 as high honour student. She started her master's degree in the Applied Linguistics Program at Karadeniz Technical University in 2021.

UZUNER is single and speaks English.

