جـامـعـة الـشـرق الأوسـط
MIDDLE EAST UNIVERSITY
Amman - Jordan

Collocational Patterns of Native and Non-Native English- Speaking Students in Theses Abstract: A Corpus-Assisted Study

المترادفات اللفظية لدى الطلبة الناطقين باللغة الإنجليزية وغير الناطقين بها في ملخصات الرسائل الجامعية: دراسة بمساعدة المتون النصية

Prepared by
Mahmoud Ali Mheize

Supervised by
Dr. Wajed Rasmi Al Ahmad

A Thesis Submitted in Partial Fulfillment of the Requirements for the Master's Degree in English Language and Literature

Department of English Language and Literature
Faculty of Arts and Educational Sciences
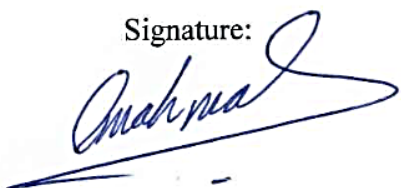**Middle East University**
May, 2024

Authorization

I, **Mahmoud Ali Mheize,** authorize the Middle East University to provide copies of my thesis on paper and electronically to libraries, organizations, and institutions concerned with scientific study and studies upon request.

Name: Mahmoud Ali Mheize

Date: 27 / 5 / 2024.

Signature:

Thesis Committee Decision

This thesis entitled "Collocational Patterns of Native and Non-Native -English-Speaking Students in Theses Abstract: A Corpus-Assisted Study" was successfully defended and approved on 27ᵗʰ / 5 / 2024.

Examination Committee Members

Name	Workplace	Title	Signature
1. Dr. Wajed Rasmi Al Ahmad	Middle East University	Supervisor	
2. Dr. Nasaybah Walid Awajan	Middle East University	Internal Examiner – Committee head	
3. Dr. Juhaina Maan Al-Issawi	Middle East University	Internal Examiner	
4. Prof. Mahmoud Ali AL-Qudah	Princess Sumaya University for Technology	External Examiner	

Acknowledgement

"Allah will exalt those who believe among you and those who were given knowledge in degrees. And Allah is knowing of all things." (Al-Mujadilah 58:11)

I am grateful to the academic English team at MEU - Dr. Nasaybah Awajan, Dr. Linda Al-Abbas, and Dr. Nisreen Yousef, for their commitment to teaching and helping students. The expertise and abilities I acquired from your guidance have been essential for this project.

I am extremely grateful to Dr. Wajed al Ahmad, who has provided me with invaluable guidance, insightful feedback, and unwavering dedication throughout this research project. Your knowledge and guidance played a key role in developing this project.

May Allah (SWT) bless all those who have contributed to this experience.

Dedication

To my cherished late Mom, (Jamilah Younis) your laughter, a radiant echo that forever warms my heart. You are forever with me in spirit.

Dear father, (Ali Mheize) your consistent support and faith in me have played a crucial role in my achievements. Your presence in my life fills me with immense gratitude.

I express my sincere gratitude to all my family and friends for their constant support and comprehension. During difficult times, your patience, good wishes, and encouraging words have been the driving force behind my perseverance.

I am extremely grateful to Dr. Wajed al Ahmad, who has provided me with invaluable guidance, insightful feedback, and unwavering dedication throughout this research project. Your knowledge and guidance played a key role in developing this project.

Table of Contents

List of Table

List of Figures

List of Appendices

List of Abbreviations

BNC	British National Corpus
ICLE	International Corpus of Learner English
EFL	English as a Foreign Language
ESL	English as a Second Language
EAP	English for Academic Purposes
ESAP	English for Specific Academic Purposes
ACLAH	Academic Collocation List for Arts and Humanities
NNS	Non-native speakers
NS	Native speakers
ESP	English for Specific Purposes
G1	Noun + Preposition
G2	Noun + To + Infinitive
G3	Noun + That-Clause
G4	Preposition + Noun
G5	Adjective + Preposition
G6	Predicate Adjective + To + Infinitive
G7	Adjective + That Clause
G8	Collocational verb patterns

Collocational Patterns of Native and Non-Native English- Speaking Students in Theses Abstract: A Corpus-Assisted Study

Prepared by

Mahmoud Ali Mheize

Supervised by

Dr. Wajed Rasmi Al Ahmad

Abstract

This study examines the collocational patterns used by native and non-native English-speaking students in theses abstracts in the fields of linguistics, literature, and translation studies with the help of a corpus. Using AntConc software, the study examines M.A. theses abstracts from Jordanian universities (Middle East University and the University of Jordan). Then the researcher contrasts them with abstracts from native English-speaking students at the OhioLINK Electronic Theses & Dissertations Center. The examination concentrates on recognizing and contrasting the occurrence, probability, and impact of grammatical collocations, particularly noun-preposition (G1) and verb-preposition (G8) structures, within the various sub-corpora. The research showed similarities and differences in collocational use among native and non-native student writers. Even though universal linguistic tendencies are indicated by shared patterns like G1 and G8, differences in frequency distribution demonstrate the impact of language proficiency, culture, and communication norms. Moreover, the study recognized specific patterns that are exclusive to the native sub-corpora, suggesting unique linguistic characteristics or stylistic preferences found in native English academic writing. The results of this research enhance knowledge of cross-cultural differences in academic writing and offer valuable perspectives for teachers and students learning languages. The research looks at the difficulties and advantages of using collocations, providing suggestions for enhancing the teaching of academic writing and helping non-native English speakers effectively present their research results.

Keywords: Collocation patterns; Native and non-native user; Corpus Driven.

المترادفات اللفظية لدى الطلبة الناطقين باللغة الإنجليزية وغير الناطقين بها في ملخصات الرسائل الجامعية: دراسة بمساعدة المتون النصية

إعداد

محمود علي مهيزع

إشراف

الدكتورة: وجد رسمي الأحمد

الملخّص

تبحث هذه الدراسة في أنماط التوافق اللفظي (Collocational Patterns) التي يستخدمها الطلاب الناطقين بالإنجليزية وغير الناطقين بها في ملخصات أطروحاتهم العلمية في مجالات اللغويات والأدب ودراسات الترجمة وذلك باستخدام مجموعة متون النصوص (Corpus) باستخدام برنامج (AntConc) يفحص البحث ملخصات الأطروحات من الجامعات الأردنية (جامعة الشرق الأوسط والجامعة الأردنية) وتقوم بمقارنة هذه الملخصات مع ملخصات الطلاب الناطقين بالإنجليزية في مركز أطروحات ورسائل الدكتوراه الإلكترونية أوهايو

(OhioLINK Electronic Theses & Dissertations Center) يركز البحث على تحديد ومقارنة حدوث واحتمالية وتأثير التراكيب المعجمية، خاصة تراكيب اسم + صفة (N-Adj) واسم + فعل (N-V) واسم + حرف جر (N-Prep) داخل مجموعات النصوص الفرعية المختلفة. يُوضح البحث أوجه التشابه والاختلاف في استخدام التراكيب المعجمية بين كتابات الطلاب الأصليين وغير الناطقين بها. وعلى الرغم من أن الاتجاهات اللغوية العالمية تتضح من خلال أنماط مشتركة مثل N-Adj و N-V وN-Prep ، فإن الاختلافات في توزيع التردد يوضح تأثير كفاءة اللغة والثقافة وقواعد الاتصال .علاوة على ذلك، تتعرف الدراسة على أنماط معينة تقتصر على مجموعات النصوص الفرعية الأصلية، مما يشير إلى خصائص لغوية فريدة أو تفضيلات أسلوبية موجودة في الكتابة الأكاديمية الإنجليزية للناطقين بها و تساهم نتائج هذا البحث في تعزيز المعرفة بالاختلافات الثقافية في الكتابة الأكاديمية وتقدم وجهات نظر قيمة للمدرسين والطلاب الذين يتعلمون اللغات .ينظر البحث في صعوبات ومزايا استخدام التراكيب المعجمية، ويقدم اقتراحات لتحسين تدريس الكتابة الأكاديمية ومساعدة الناطقين بغير الإنجليزية على عرض نتائج أبحاثهم بشكل فعال.

الكلمات المفتاحية: المترادفات اللفظية، الناطقين باللغة الإنجليزية وغير الناطقين بها، دراسة مبنية على متون النصوص.

CHAPTER ONE
Introduction

1.1. Background of the Study

A thesis abstract serves as the initial point of reference for individuals interested in investigating a research study, as it offers a concise summary of the research's objectives, methodologies, findings and conclusions. It functions as an entryway, aiding readers in making informed decisions regarding whether they should allocate their time to perusing the complete research. This choice significantly influences the extent to which research findings are published. Within the academic research scope, time is a valuable and limited resource, thus, abstracts play a vital role in simplifying decision-making processes. They enable readers to efficiently comprehend the principal aspects of the research, thus facilitating the prioritization of their reading selections.

The difficulties of employing collocations in academic writing, particularly in research papers, prompted the conduct of this study. This study examines abstracts of papers authored by both native and non-native students.

Theses abstracts are prominently featured in academic databases, search engine results, conference programs, and various platforms, making them the first pieces of content that reach a broad audience. Their prominence underscore their significant role in determining the accessibility and distribution of the research. The quality of a thesis abstract is like the basis for the entire research. A carefully written abstract usually indicates a well-structured, thoroughly researched, and rigorously conducted research paper. On the other hand, a poorly written abstract can discourage potential readers, potentially reducing the research's visibility and influence. This highlights the vital role that the abstract plays in shaping how the research is received and shared.

The examination of abstracts written by native and non-native students, allow the exploration of potential differences in language and culture when it comes to composing abstracts. This is especially relevant given the increasingly diverse academic landscape, where certain academic fields may be significantly influenced by non-native speakers of a particular language. By considering factors like language proficiency and cultural influences in the creation of abstracts, we can improve the clarity and accessibility of research. For educators and academic institutions, the assessment of abstracts written by non-native students provides valuable insights into the specific challenges these students face when writing academically. This understanding can guide teaching methods, enabling tailored support and instruction in the intricacies of academic writing. Ultimately, it enhances the educational experience for non-native students. For example, it is possible that non-native student writers may be more likely to use certain collocations that are less common in academic English, or they may use collocations in a different way.

1.2. Statement of the Problem

This study aims to address the research gaps concerning the use of grammatical collocations in abstracts written by Jordanian students in disciplines such as linguistics, literature, and translation studies in comparison to abstracts produced by native student writers. No research addressed the existing literature no studies examined the collocational patterns employed by Jordanian students in these academic areas compared to those of their native peers. Consequently, this research attempts to fill this gap by providing insights into the writing practices of Jordanian students within these academic domains. Through this contribution, the study aims to enhance the understanding of cross-cultural variations in academic abstracts. in addition, This study

was prompted by the difficulties second language learners encounter in understanding collocations and their meanings. In order to promote better comprehension and use of grammatical collocations, a comprehensive study of these collocations was conducted.

1.3. Objectives of the Study

The primary objective of this study is to conduct a comparative study of collocational patterns in theses abstracts written by both native and non-native student writers in linguistics, literature, and translation studies. The analysis covers grammatical collocations and aims to shed light on similarities and differences in relation to collocational pattern use in academic writing. By exploring language choices, the research seeks to enhance language awareness, particularly for non-native students, and provide pedagogical insights for educators. The study employs a corpus-driven approach, using large collections of naturally occurring texts for empirical analysis.

1.4. Questions of the Study

The current study answers the following questions:

1. What are the most common grammatical collocational patterns used by native and non-native students when composing theses abstracts?

2. What are the similarities and differences between the use of collocations by native and non-native student writers across genres (linguistics, literature, and translation studies) in terms of frequency?

1.5. Significance of the Study

The analysis of theses abstracts, especially when considering native and non-native student writers, holds profound significance in the academic writing scope. It examines grammatical collocations, which offers insights into the writing practices of both native

and non-native student writers. This analysis contributes to the educational experience of non-native students related to academic research writing and pedagogy, by Identifying Common Errors, by comparing the collocations used by native and non-native students, the analysis can identify areas where non-native students may be making errors. For example, it may identify excessive usage of verb-noun pairs or prepositional phrases. This information might be used to develop specialised learning materials that target these specific challenges.

Exposing Students to Natural Language, the analysis can provide non-native students real-world examples of how native speakers use collocations in academic writing. This exposure can provide students with a better knowledge of natural and idiomatic phrasing, leading in more sophisticated and nuanced writing.

Informing Pedagogical techniques, the study's findings can be used to develop instructional techniques that focus on the use of collocations in academic writing, For example, instructors may include activities that focus on finding and practicing frequent collocations across academic fields.

CHAPTER TWO
Review of related literature

2.1. Collocations

Collocations are fundamental linguistic units falling within the broader category of chunks, coined by John Firth (1957) as words to remain a companion. They are basically word combinations where the words they associate with have an impact on their individual meanings. As defined by Sinclair in 1966, a collocation is made up of a node word which serves as the core lexical unit, and the "collocate", which is the closely linked word to it. These lexical combinations form the basis of corpus linguistics, offering significant insights into the complex structure of natural language use. It has been argued that collocations refer to words that regularly appear together in a text, emphasizing their strong connection and close placement (Sinclair, 1991). However, it has been reported that collocations do not always have to be adjacent; they can be separated by other words in a sentence or passage. This flexibility in collocational structures enriches our comprehension of the dynamics of language (Nesselauf, 2003).

Collocations are divided into two main categories: 1- Lexical collocations and 2- grammatical collocations. The combinations of two or more content words that create distinct syntactic structures with varied degrees of meaning transparency are known as lexical collocations. The level of semantic restriction in these lexical elements is typically determined by the extent to which they can be substituted or taken literally, as seen from a phraseological perspective (Mirsalari 2019; Seretan,2005). Lexical collocations frequently include content words such as adjectives, adverbs, verbs, and nouns. These collocations involve words that naturally co-occur due to their shared

semantic relationship. For example, considering the phrase "catch a cold", the verb "catch" and the noun "cold" make sense together to suggest being sick (Chang, 2018; Demir, 2017; Farghal and Al-Hamly,2007).

On the other hand, Grammatical collocations encompass combinations of words, including nouns, adjectives, or verbs, coupled with particles like prepositions, adverbs, or grammatical structures such as infinitives, gerunds, or clauses. These combinations, often referred to as chunks or formulaic sequences, are distinctive for their fixed, almost idiomatic nature. They are commonly used in language and serve as cohesive units. Examples of these grammatical collocations include "at night," "extend to," "good at," "fall for," and "to be afraid that". While some of these expressions exhibit a strong sense of belonging together, others, such as "from the outside" or "inside the cupboard," are less tightly associated, resembling free combinations (Bahns, 1993; Moehkardi & Moehkardi, 2002). Furthermore, according to Benson's et al. (1986), grammatical collocations can be divided into multiple categories. Table (1) illustrates the different types of grammatical collocations with examples.

Table (1): BBI Classification of Collocations- Grammatical Collocations (adopted from Futagi et al. 2008, p. 5)

Label	Syntactic Pattern	Examples
G1	**Noun + Preposition**	blockade against, apathy towards
G2	**Noun + To + Infinitive**	They had the foresight to do it
G3	**Noun + That-Clause**	We reached an agreement that she would represent us incourt
G4	**Preposition + Noun**	by accident, in advance, in agony
G5	**Adjective + Preposition**	They are angry at the children
G6	**Predicate Adjective + To + Infinitive**	It was necessary to work
G7	**Adjective + That Clause**	She was afraid that she would fail her examination
G8	Collocational verb patterns:	
	Shift of an indirect object to a position before the directobject of transitive verbs is allowed.	He sent the book to his brother

Label	Syntactic Pattern	Examples
	B. Shift of an indirect object to a position before the directobject by deleting to is not allowed.	They described the book to her
	C. Transitive verb with preposition for allows the deletion offor and the shift of the indirect object to a position before thedirect object. For example:	She bought a shirt for her husband
	D. The verb forms a collocation with a specific preposition and an object.	They based their conclusions on the available facts
	E. Verbs are followed by to + infinitive.	They began to speak
	F. Verbs are followed by infinitive without to These verbs,except, dare, help, and need, are called modals.	They must work
	G. Verbs are followed by second verb in ing.	They kept talking
	H. Transitive verbs are followed by an object and to + infini-tive.	They asked the students to participate in discussion
	I. Transitive verbs are followed by a direct object and an in-finitive without to. Most I-pattern verbs cannot be passivized.	We let them use the car
	J. Verbs are followed by an object and a verb in –ing.	I caught him smoking in his bedroom
	K. Verbs can be followed by a noun or pronoun and gerund.	This fact justifies Bill's coming late They love his clown-ing
	L. Verbs are followed by a noun clause beginning with con-junction that.	They admitted that they were wrong
	M. Transitive verb can be followed by a direct object, an in-finitive to be and adjective / past participle/ noun/pronoun.	We considered her to be very capable, well-trained, acompetent engineer
	N. Transitive verbs are followed by a direct object and adjec-tive/ past participle or noun/pronoun.	She dyed her hair red
	O. Transitive verbs are followed by two objects.	The teacher asked the students questions
	P. Intransitive /reflexive / transitive verbs must be followed by an adverbial (an adverb/a prepositional phrase/a noun phrase/a clause).	He carried himself with dignity; but not * He carried him-self
	Q. Verb can be followed by an interrogative word, such ashow, what, when, etc	He always wants what I want She knows when to keepquiet

Label	Syntactic Pattern	Examples
	R. Dummy 'it' is followed by transitive verbs (often express- ing emotions) and by to + infinitive or by that + clause or byeither.	It puzzled me that he never answered the telephones
	S. A small number of intransitive verbs are followed by apredicate noun / predicate adjective. Including the verb make, used intransitively, belongs to this group.	She will make a good teacher

To conclude, collocations provide a detailed understanding of how words function together in language and are a fundamental component of both general and corpus linguistics. They can be characterized as collections of words that co-occur frequently because of their relationships in grammar or semantics. Lexical and grammatical collocations are the two primary forms of collocations that add to the richness and depth of language. The historical significance of collocations in linguistics emphasizes their enduring importance in the study of language structure and meaning. By examining collocations in a corpus, linguists gain valuable insights into the complex patterns of word usage that underlie effective communication.

Furthermore, understanding collocations is crucial for academic writing, as students with brilliant ideas often lose points due to their lack of familiarity with common word combinations. In the case of second language (L2) students, their writing frequently becomes weaker because of this deficiency. For instance, an essay on smoking included the phrase "the smokers who rely on cigarettes and have to smoke every day" instead of employing more appropriate expressions like "heavy smoker" or "addicted to smoking." (Zou & Thomas, 2018). Thus, learning collocations hold particular importance for L2 learners, facilitating improvements in both spoken and written language.

2.2. Corpus of Linguistics

Corpus linguistics is the study of language using large collections of text, or corpora. Corpus linguists use computer-aided tools to analyze and process these corpora. Corpus linguistics allows researchers to access and analyze vast amounts of text, which can be used to study a variety of linguistic phenomena, such as word usage, grammar patterns, and language change. Tognini-Bonelli (2001) first introduced the distinction between corpus-based and corpus-driven language research. Corpus-based research uses corpus data to test or validate existing theories or hypotheses, while corpus-driven research uses corpus data to generate new theories and hypotheses about language. Basically, corpus-based research is a logical methodology that begins with a theory or hypothesis and tests it using data from a corpus. Using corpus data as a starting point, corpus-driven research is an inductive methodology that produces new theories or hypotheses. The current analysis is a corpus- driven analysis using data to come out with hypotheses about the use of grammatical collocations by Jordanian graduate students' writings and native graduate students. With the advancement of computer technology, using corpora has become easier. Four primary categories of corpora exist, comprising:

1. Generalized corpora: These collections of text include a broad variety of content from many publications, including books, periodicals, newspapers, and websites.

2. Specialized corpora: These collections of data concentrate on a particular subject or field, including academic, legal, or professional language.

3. Comparable corpora: These corpora contain texts in two or more languages that are similar in terms of topic, genre, and register.

4. Parallel corpora: These collections of texts comprise translations of one another into two or more languages.

Specialized corpora contain text from a particular domain or topic, whereas generalized corpora contain text from a variety of sources. Researchers can draw generalizations about language use because generalized corpora are usually significantly bigger than specialized corpora. Conversely, specialized corpora can be employed to research particular linguistic traits or usage patterns. A generalized corpus example would be: More than 100 million words of text from a range of sources, including books, periodicals, newspapers, and websites, which make up the British National Corpus (BNC).

Whereas specialized corpora are compiled in accordance with specific research purposes. Learner corpora are considered one of the specialized corpora. They are collections of text written by language learners. They are employed to investigate the linguistic characteristics of language learners that set them apart from native speakers. The International Corpus of Learner English (ICLE), which has over 20,000 words of text from learners in over 75 nationalities, is one example of a learner corpus. Another type of specialized corpora is pedagogic corpora which are collections of classroom vocabulary and other languages that learners have been exposed to. They are used to see if learners are being taught useful language and to help teachers improve their skills. Both learner corpora and pedagogic corpora can be valuable resources for language teachers and researchers. Learner corpora can be utilized to pinpoint areas in which students require additional assistance, and pedagogical corpora can guarantee that students are receiving instruction in the language they require conversely, collections of texts in two or more languages that are translations of the same original text are known as parallel corpora.

This means that parallel corpora contain identical texts in both their original and translated versions. Parallel corpora can be used to compare linguistic and discourse patterns between languages without introducing translational errors. Comparable corpora, on the other hand, are collections of text in two or more languages, where the texts are similar but not necessarily translations of each other. This implies that similar corpora might include texts that are not exact translations of writings on the same subject, genre, or register. Numerous linguistic phenomena, including cross-linguistic diversity, language change, and translation techniques, can be studied with comparable corpora (Lei & Liu, 2018).

The current study analyzes collocations through specialized corpora, which are compiled to answer the specific research questions and a data- driven analysis that highlight how native and non-native student writers employ collocations in their abstracts. The corpora are collected and compiled for the purposes of the study to answer the research question related the similarities and differences between the use of native and non- native students .

2.3 Empirical Studies related to

This section provides studies conducted analyzing the use of collocations by native

A corpus-based study by Durrant & Schmitt (2009) analyzed the use of collocations by native and non-native English writers, specifically looking at common ones. Analysts examined 96 texts in different styles (essays, research papers, articles, compositions) authored by individuals who were both native and non-native speakers. They concentrated on noun phrases that are right next to each other (such as adjective-noun, noun-noun). The results revealed that native speakers utilized a broader variety of uncommon word pairings in comparison to non-native speakers. Both groups utilized powerful combinations of words (with high T-scores) at comparable frequencies. Nevertheless, non-native writers had a tendency to excessively utilize specific preferred powerful word combinations. Non-native writers show a significant tendency to use collocations with lower mutual information scores (weaker association) compared to native writers. The gap was even larger when repetition was not taken into account, indicating a limited range of these less common but crucial word combinations. The findings indicated that non-native writers tend to choose well-known, common collocations over rarer but more authentic ones. Although non-native speakers can easily learn common collocations, they face difficulties when trying to grasp less common but highly connected phrases that are crucial for achieving fluency comparable to that of native speakers. This endorses a language learning model that is based on usage, where the frequency of exposure plays a vital role in acquiring formulaic language. Language instructors may have to specifically concentrate on high MI collocations to assist students in enhancing their comprehension of natural language patterns.

and non-native students. Ramos, & González (2013) compared the collocational richness of learners and native speakers based on four parameters: number of collocations relative to the total number of words, number of different collocations used, use of low-frequency collocations, and quantity of incorrect collocation usage. Learners do use collocations, but they demonstrate less variety, sophistication, and accuracy compared to native speakers. The study challenged the assumption that learners rarely use collocations and highlights the importance of focusing on collocational competence in language learning. Error analysis revealed distinct patterns in learner errors (often due to interlingual interference) and native speaker errors (primarily intralingual interference). The findings suggested a need for increased emphasis on collocation instruction in language teaching materials and methods, including developing materials that address collocations, integrating error analysis and correction strategies, and encouraging learners to explore and experiment with collocations.

On the other hand, Dukali (2016) investigated the challenges faced by Libyan undergraduate English major students in the use of verb-noun and adjective-noun collocations through an examination of their performance in free writing. The study focused on twelve verbs and twelve adjectives to determine how Libyan learners employed these collocations and assessed whether native English speakers perceive differences in the acceptability of learner-generated collocations in academic versus non-academic contexts. To achieve this goal, the study gathered data from fourth-year university students (90 males and 96 females) at Tripoli University, specifically from the Department of English within the Faculty of Arts. A 250-word academic writing task served as the primary data collection method.

The data analysis employed AntConc 3.2.1w (Anthony, 2007), and learner-generated collocations were assessed for acceptability using four sources: the Oxford Collocations Dictionary (2009), the online British National Corpus (BNC), consultations with two native speakers, and a survey involving 100 native speakers of English to triangulate the results. The analysis framework of Gass and Selinker (2008) for error analysis was used to assess learners' collocational violations, with both quantitative and qualitative methods applied to the data. Additionally, questionnaires were administered to students and lecturers (155 students and 12 university lecturers) to gain insights into the teaching and learning of collocations in the Libyan educational system. The research findings revealed that verb-noun collocations posed more significant challenges to the participants than adjective-noun collocations. Various types of errors were identified, including grammatical, lexical, and usage-related errors. Eight main sources of difficulties were suggested, such as mother tongue interference and synonym usage. The survey data indicated significant differences in native speakers' judgments between academic and non-academic contexts. Based on these results, the study offers recommendations to enhance the teaching of collocations in EFL (English as a Foreign Language) classes, addressing the challenges identified in the research.

Moreover, Farooqui (2016) examined the use of collocations among both non-native speakers (NNS) and native speakers (NS) of English in the context of Computer Science academic writing. He compared these language patterns in Master's dissertations written by NNS and NS students with those found in research articles by expert writers. The study aimed to understand whether NNS students overused or underused certain collocations, and if so, why. They also developed activities to help NNS students improve their collocation use. The researcher collected and analyzed data

from 55 student dissertations and 63 research articles, focusing on the most common noun and verb collocations. The findings showed that both NNS and NS students tended to overuse noun collocations, while verb collocations were not significantly different from the reference corpus. The overuse of noun collocations was attributed to factors like sub-discipline, topic, and genre. The study suggested that teaching collocations in specific academic contexts, considering expertness, genre, and discipline-specific factors, can be beneficial for English for Specific Purposes (ESP) instruction and has implications for Applied Linguistics and teaching practices.

Thongvitit and Thumawongsa also (2017) investigated the usage patterns and frequency of both grammatical and lexical English collocations within research article abstracts in the liberal arts and humanities, authored by Thai EFL writers between 2010 and 2015. The research utilized software tools, specifically Antconc and TagAnt, for data analysis, and the Oxford Collocations Dictionary for Students of English (second edition, 2009) to identify misused collocations. The primary findings indicated that noun + preposition and adjective + noun collocations were the most frequently employed grammatical and lexical collocations in these abstracts. Additionally, the study identified noun + preposition and verb + noun collocations as the most misused ones. These results offer insights into the use of English collocations by Thai EFL writers in the context of liberal arts and humanities research articles, potentially informing language instruction and academic writing support for this specific group of writers.

In their study, Mousavi and Darani (2018) addressed a shift in corpus linguistics research toward investigating word co-occurrence rather than individual words, with a particular focus on recurrent word combinations in academic prose. The study examined

the most frequent collocations used by Iranian and Norwegian scholars in a corpus of 17 articles from the Journal of Pragmatics, using a frequency-based approach. The corpus comprised nine articles by Iranian scholars (67,673 words) and eight articles by Norwegian scholars (64,682 words), with data collected using Collocation Extract software. The results were presented in three phases. In the first phase, it identified the 15 most frequent lexical collocations in both corpora, classified into three types of lexical collocations. The "Adj+N" collocation type had the highest proportion in the corpora, while "Adv+Adj" had the lowest proportion. In the second phase, the study presented the lexical collocations within the Iranian corpus, identifying a total of 818 collocations categorized into five types. "Adj+N" was the most frequent, and "N+V" was the least frequent. Similarly, the lexical collocations in the Norwegian corpus were categorized into four types, with "Adj+N" being the most frequent and "Adv+Adj" the least frequent. In the third phase, the study compared the frequencies of lexical collocations in the two corpora and found no significant difference in the use of all collocation types, except for "Adj+N," where there was a notable distinction.

O'Flynn (2018) also analyzed collocation, attempting to study English for academic purposes (EAP) by examining the commonly used lists of academic vocabulary, such as the Academic Word List (Coxhead, 2000), academic collocation list (Ackerman and Chen, 2013), and academic formulas list (Simpson-Vlach and Ellis, 2010). This analysis set out to establish the necessity for an empirically derived discipline-specific list of academic collocations, particularly in the context of English for Specific Academic Purposes (ESAP). It attempted to evaluate the Academic Collocation List for Arts and Humanities (ACLAH). The findings indicated that ACLAH represents a significant step toward a more comprehensive understanding of academic collocations, which can better

assist EAP students in Arts and Humanities compared to generic academic vocabulary lists. The study concluded by outlining implications for future research and teaching methods involving ACLAH.

Using qualitative approach, Sari & Gulö (2019) analyzed students' essays to identify grammatical collocation errors in English language learners' writing. They examined 15 essays from first-year English Literature students, classifying errors based on eight grammatical collocation types. Dictionaries served as references to identify correct and incorrect collocations. The analysis revealed 35 errors, with the most frequent involving prepositions (A collocations) and verbs (V collocations). The findings suggest learners struggle with prepositions within collocations, possibly due to native language interference. The study highlights the importance of explicitly teaching collocation rules and emphasizes incorporating targeted instruction on grammatical collocations into the curriculum, including activities that compare collocations in English to students' native language and tasks focused on identifying and correcting errors in authentic texts.

Further, Duong & Nguyen (2021) examined the role of collocations in improving academic writing for EFL (English as a Foreign Language) learners, specifically focusing on students and lecturers at Van Lang University in Vietnam. The study involved 47 English-major juniors and 5 English lecturers. Two primary instruments were used: Online questionnaires for students to gauge their perception and usage of collocations. Email interviews with lecturers to gather their insights on student challenges and teaching approaches related to collocations. Descriptive statistics were employed to analyze quantitative data from the questionnaires. Thematic analysis was used to identify key themes and patterns within the qualitative data from both

questionnaires and interviews. The Findings shows the Benefits of Collocations. Both students and lecturers acknowledged the significant role of collocations in enhancing academic writing. The study underscores the need for explicit teaching of collocations in EFL classrooms. Teachers should actively highlight collocations within authentic texts and provide guidance on their usage. English to students' native language and tasks focused on identifying and correcting errors in authentic texts.

Another study by Wu and Halim (2022), who conducted Corpus-based on L2 collocational knowledge this study examined existing research on how learners of English as a Second Language (ESL) and English as a Foreign Language (EFL) acquire knowledge of collocations (word combinations with specific meanings). Most research focused on learners' ability to use collocations (productive knowledge) rather than understanding them (receptive knowledge).

With a particular focus on grammatical collocations in the domains of linguistics, literature, and translation studies, this study explores the collocational patterns used in theses abstracts by native and non-native English-speaking students. Using a corpus-driven methodology, the study compares abstracts from native English-speaking students at the OhioLINK Electronic Theses & Dissertations Centre with a set of abstracts from Jordanian universities, Middle East University and the University of Jordan. The prevalence, probability, and impact of grammatical collocations—specifically, noun-preposition (G1) and verb-preposition (G8) structures—are examined in this study. The results show that native and non-native writers use collocations in different ways and in comparable ways. Variations in frequency distribution show the impact of language competency, cultural context, and communication, even when common patterns like G1 and G8 are shared.

This study supports the notion that language competency influences collocational usage by identifying a difference in frequency distribution between native and non-native writers, which is similar to prior corpus-based studies on collocational patterns. Similar to Durrant & Schmitt (2009), the study demonstrates that non-native authors could overuse a few popular collocations, while native speakers often use a wider variety of collocations. This study, however, departs from previous research in that it examines the impact of academic field on collocational patterns and discovers that the distribution of patterns varies between linguistics, translation studies, and literature. Farooqui (2016), who also looked at discipline-specific determinants on collocation utilization, is in accord with this. The research also finds distinct patterns that are only present in native corpora, which is consistent with Lee & Swales' (2006) findings.

CHAPTER THREE
Methodology

3.1. Corpus Compilation

The current study will analyze the collocational patterns used by native and non-native writers of theses abstracts written by non-native student writers (Middle East University -MEU & University of Jordan- UJ) and theses abstracts written by native student writers from OhioLINK- Electronic Theses & Dissertations Center (https://etd.ohiolink.edu/acprod/odb_etd/r/etd/search/1?clear=0,1,5,10,20,21,1001).

The criteria for choosing the abstracts were the accessibility of them, since the researcher was able to access the library of MEU and University of Jordan.

Each divided into three main categories: linguistics, literature, and translation studies. Each category (sub-corpus) comprised 20 abstracts with 120 abstracts in both native and non-native corpora. The size of the native corpus is 13452 words while the size of the non- native corpus is 13560 words. Then, the texts were uploaded into TagAnt and AntConc for analysis after converting them into a TXT file.

3.2. Instrument of the Study

1-AntConc

AntConc, a user-friendly, free software tool developed by Laurence Anthony, a professor of applied linguistics at Waseda University in Tokyo, Japan, serves as a crucial resource for corpus linguistic analysis. This software is widely employed by linguists, researchers, and language educators to dissect textual data and unveil language usage patterns and trends. It provides a versatile range of features for exploring language patterns and frequencies in extensive datasets. Users can import text

files or entire folders of text files into AntConc to create a corpus, which can then be subjected to various analytical tools and techniques (Froehlich, 2015). These tools encompass:

1. Concordance: This feature allows users to search for specific words or phrases in the corpus and generates a concordance, displaying each instance in context. It helps to identify language use patterns like collocations and word associations.

2. Collocation: AntConc includes a collocation tool that identifies frequently occurring word pairs in the corpus, helping users recognize lexical patterns and understand word combinations in specific contexts.

3. Cluster/Cluster analysis: This tool identifies groups of words that frequently co-occur in the corpus, assisting in the identification of themes or topics within the data.

4. Word list: This tool generates a list of unique words in the corpus, along with their frequency of occurrence, facilitating the identification of the most used words.

This study employed wordlists, concordance and collocation functions to study the collocational patterns beside the employment of TagAnt software.

2-TagAnt

Laurence Anthony created TagAnt, a free software program specifically made for corpus linguistics projects. It serves as a segmenter for multiple languages and also performs part-of-speech tagging (POS) . Part-of-speech tagging involves giving grammatical labels (such as noun, verb, adjective) to every word in a text . This data is essential for a variety of uses in computational linguistics, such as sentiment analysis, machine translation, and information retrieval.

TagAnt uses the TreeTagger software engine, developed by Helmut Schmid . This tool enables TagAnt to manage different languages, making it a flexible tool for researchers who work with multilingual datasets. The software provides a friendly interface allowing users to enter text or a list of text files for POS tagging. The result displays the initial content accompanied by the designated POS labels for every term (Jurafsky, & Martin, 2009).

Even though TagAnt is a valuable tool for researchers, it is crucial to recognize certain constraints. The software can be used for individual and non-profit research for free, but group use requires permission. Moreover, the tagset used by TagAnt (Penn Treebank) is not the same as the one utilized in other commonly used corpus analysis tools. This difference requires caution when comparing TagAnt results with those of other software programs.

To sum up, TagAnt offers a convenient and successful tool for POS tagging in corpus linguistics studies. Researchers find great value in the asset due to its ability to work with various languages and be used for non-commercial purposes. Nevertheless, users must take into consideration licensing limitations and possible inconsistencies in tagsets when incorporating TagAnt into their process.

3.3. Data analysis

3.3.1. Language Patterns: Likelihood and Effect in Corpus Linguistics.

Corpus linguistics relies greatly on statistical methods to study language using large text collections (corpora). ways to discover patterns and significance. Two important ideas in this examination are probability and impact, which function Working together to create a more detailed depiction of word usage.

Probability, commonly evaluated using metrics such as log-likelihood, determines the chances of a certain word appearing in a specific situation versus a haphazard arrangement [McEnery & Hardie, 2012]. Pretend to explore the term Comparison between "happiness" in news articles and social media posts. A high log-likelihood score for the word "happiness" indicates...social media seems to show up more often than would be expected by random chance, indicating a possible emphasis on emotions. On that particular platform. On the flip side, impact explores further. It delves deeper than just the frequency of a word and investigates the extent of the variation in its usage across various contexts.

This can be quantified using figures and data such as probabilities. proportion. Expanding on the example of "happiness," an odds ratio analysis could reveal a tenfold rise in "happiness" in. Social media can be contrasted with news articles. This measures the important impact that social media has on the communication of. this idea. By using both probability and impact, corpus linguists can advance beyond simple frequencies. They have the ability to not just determine words that are statistically significant and assess the magnitude of these variances. This unified method enables researchers to make more detailed assertions regarding language usage. An example would be corpus linguists studying the probability and impact of euphemisms related to death in medical contexts. differences between journals and obituaries (Stubbs, 2008). There is a high probability of euphemisms present in both datasets, but the impact is uncertain.

The strength of the association (as indicated by odds ratio) could be greater in medical publications, indicating a more objective view towards mortality. Ultimately, probability and impact are essential instruments.

3.3.2. Frequency analysis in Corpus Linguistics

Frequency analysis within corpus linguistics entails the quantification of word or phrase occurrences in a corpus, which is essentially an extensive compilation of written materials or text. This analytical approach yields valuable insights into the significance of linguistic elements and their usage patterns within a specific corpus. Two primary components of frequency analysis are raw frequency and normalized frequency (Hardie,2015).

3.3.3. Raw Frequency

Raw frequency involves a straightforward tally of how often a particular word or phrase appears within a corpus. It provides a basic measure of the frequency of a specific linguistic unit in the given text. For instance, if the word "apple" appears 50 times in a corpus containing 10,000 words, its raw frequency is 50.

3.3.4. Normalized Frequency

However, normalized frequency is a more advanced statistic that takes the size of the corpus into account. It can be calculated by dividing the raw frequency of a word or phrase by the entire number of words in the corpus, and then scaling the result by a constant factor, typically 100, 1,000, or 10,000. Even though the sizes of the corpora differ, word frequencies from different corpora can be compared thanks to normalization.

Significance of Normalized Frequency Analysis:

Normalized frequency analysis holds considerable importance within the field of corpus linguistics for several compelling reasons:1). Cross-Corpus Comparison: Diverse corpora can vary significantly in size, encompassing both small, specialized collections and large datasets. Normalized frequency permits researchers to impartially compare the

frequencies of words or phrases across these varied corpora. This is particularly vital for identifying language usage patterns and variations, such as comparing word frequencies between academic texts and casual conversations. 2). Assessment of Relative Importance: Normalized frequency offers a more precise assessment of the relative significance of words or phrases within a corpus. It aids researchers in distinguishing between high-frequency terms that are commonly used in the language as a whole and those that hold particular significance within a specific context or corpus.   3). Identification of Key Terms: Researchers frequently employ normalized frequency analysis to pinpoint key terms or expressions of specific interest within a corpus. By examining which words have the highest normalized frequency, they can identify crucial concepts, themes, or specialized terminology within the text.4). Corpus Profiling: Normalized frequency analysis is a fundamental tool for corpus profiling, which involves generating a comprehensive characterization of a corpus's linguistic attributes. Researchers can harness normalized frequency to identify the most distinctive words and phrases that typify a particular corpus, thus enhancing linguistic and textual analysis. Thus, normalized frequency analysis is a pivotal technique in corpus linguistics that facilitates equitable comparisons of word and phrase frequencies across corpora of varying sizes. It furnishes insights into the relative importance of linguistic elements within a specific context and supports an array of research applications, including the identification of key terms, corpus profiling, and the exploration of language variations. Corpus tools can be used to measure the strength of collocations in several ways (Al Ahmad, 2021; Gries, 2009; Hardie,2015).

3.4. Procedures of the study

1- Reviewing theoretical literature related to collocational analysis and corpus analysis.

   By examining hypotheses on the co-occurrence of words and the implications for meaning and language use. Using corpus analysis particularly collocational analysis

2- Reviewing empirical studies relevant to collocational analysis and corpus analysis.

   Search for academic research written by other researcher's proceedings, and other publications that demonstrate how collocational analysis and corpus analysis have been applied in various contexts.

3- Selecting the abstracts for the study and conducting initial analysis of collocational patterns. For native and non-native writers of theses abstracts written by non-native student writers (Middle East University -MEU & University of Jordan- UJ) and theses abstracts written by native student writers from OhioLINK- Electronic Theses & Dissertations Center

   (https://etd.ohiolink.edu/acprod/odb_etd/r/etd/search/1?clear=0,1,5,10,20,21,1001).

4- Presenting and reporting the findings along with discussing them in light of previous research. The study analyzed collocational patterns in thesis abstracts from native and non-native English-speaking students in linguistics, literature, and translation studies. It found similarities and differences in grammatical collocations, with both groups relying on G1 and G8 structures. However, native writers had a wider range of patterns and unique patterns, suggesting linguistic nuances specific to native English writing.

5- Listing references as per the APA style.

CHAPTER FOUR
Results

4.1. Introduction

This chapter introduces the analysis of results related to the usage of collocation patterns by native and non- native writers of theses abstracts. It provides in depth analysis of grammatical collocation patterns used in six sub-corpora: native literature, native linguistics, native translation, non-native literature, non-native linguistics and non-native translation. Then, a comparison between the six sub-corpora is presented to highlight the similarities and differences between native and non-native use.

4.2. Results of Question One: "What are the most common grammatical collocational patterns used by native and non-native student writers when composing theses abstracts?"

To answer this question, the researcher used TagAnt software to find the collocate patterns by determining the grammatical category of words (noun, verb, adjective, adverb, preposition), then uploading the tagged files into AntConc software to examine the collocates that occur with the node words, with a focus on likelihood and effect. Likelihood measures the probability of a specific word appearing next to another word (collocate) compared to random chance. The effect, on the other hand, considers likelihood but adjusts for the overall frequency of the individual words. A positive effect suggests the collocation appears more frequently than expected, while a negative effect indicates a less frequent occurrence than expected. These two features are provided by AntConc to measure the collocation association strength, which facilitates the search for collocations.

4.2.1. Native sub-corpora:

The table below demonstrates the grammatical patterns used in Native literature sub-corpus. The tags used are: noun (NN*), verb (VB*), adjective (JJ), preposition (IN), whereas the categories are divided in accordance with Benson et al, 2009 classification of grammatical collocations.

Table (2). Grammatical collocation patterns in Native- literature sub-corpus

Collocational pattern	Frequency	likelihood	Effect
G1(NN*_IN)			
NN*_of	87	241.812	2.946
NN*_for	8	18.337	2.709
NN*_within	4	16.318	3.809
G2 (NN_ to infinitive)	44	41.980 1.530	1.530
G3 (NN_ that)	23	32.821	1.936
G4 (IN_ NN*)	157	447.482	2.780
G5 (JJ_ IN)	17	33.765	-1.581
G8 (VB*_ IN)			
VB*_in	8	86.047	-2.990
VB*_ of	3	21.412	-2.627
VB*_with	9	20.999	2.671

This table demonstrates the following:

1. The collocational patterns, within Native literature sub-corpus, are used with a focus on nouns (NN*), verbs (VB*), adjectives (JJ*), and prepositions (IN_). The most significant collocational pattern is G1(NN*_IN), where nouns are followed by prepositions. Specifically, nouns followed by "of" (NN*_of) such as importance of and significance of, and "for" (NN*_for), such as background for, have a high frequency and likelihood, indicating a strong tendency for these collocations to occur. Additionally, nouns followed by "within" (NN*_within),

such as working within, show a high effect, suggesting a significant impact on the meaning or structure of the text.

2. G2 (NN_to infinitive) pattern reveals a moderate frequency and likelihood of nouns being followed by "to" and an infinitive verb, indicating a preference for this collocational structure, for instance, in order to. The G3 (NN_ that) pattern shows a moderate frequency and likelihood of nouns being followed by "that", suggesting a tendency for this collocational structure to occur.

3. G4 (IN_ NN*) pattern reveals a high frequency and likelihood of prepositions being followed by nouns, indicating a strong tendency for this collocational structure to occur. The G5 (JJ_ IN) pattern, however, shows a negative effect, suggesting that adjectives followed by prepositions are less common or have a weaker impact on the meaning or structure of the text.

4. G8 (VB*_ IN) pattern reveals a mixed effect of verbs followed by prepositions. While verbs followed by "in" (VB*_in) and "with" (VB*_with) show a moderate frequency and likelihood, they have a negative and positive effect, respectively. This suggests that verbs followed by "in" may have a weaker impact on the meaning or structure of the text, while verbs followed by "with" may have a stronger impact. Verbs followed by "of" (VB*_ of), on the other hand, have a low frequency and likelihood, indicating a weaker tendency for this collocational structure to occur.

Table 2 examines grammatical collocations in native speakers' writing, revealing a strong preference for nouns followed by prepositions, particularly "of" and "for". The effect of "within" on sentence structure is high. Moderate preference is found for noun phrases followed by "to" and an infinitive verb, and nouns followed by "that".

Adjectives followed by prepositions have a negative effect. The linguistics sub-corpus's preponderance of noun phrases and particular prepositions points to a writing style that emphasises exact idea connections, definitions that are easy to understand, and technical language that is concentrated on examining how language operates.

Table (3). Examples on Grammatical collocation patterns in Native- literature sub-corpus

Label	Syntactic Pattern	Examples
G1(NN*_IN)		
G1	NN*_of	The history of developmental education
	NN*_for	a blueprint for application
	NN*_within	writers working within metropolitan
G2	(NN_ to infinitive)	The aim to investigate
G3	(NN_ that)	The assertion that aspects
G4	(IN_ NN*)	in response to the construction.
G5	(JJ_ IN)	confident about their understanding
G8 (VB*_ IN)		
G8	VB*_in	engaged in religious controversy
G8	VB*_ of	robs us of the richness
G8	VB*_with	interacted with spatial dislocations

Table (4). Grammatical collocation patterns in Native- Linguistics sub-corpus

Collocational pattern	Frequency	Likelihood	Effect
G1 (NN*_IN)			
NN*_of	166	543.004	2.532
NN*_in	69	28.169	0.751
NN*_for	16	27.224	1.924
NN*_by	16	23.030	1.766
NN*_with	15	21.821	1.776
G3 (NN_ that)	22	19.966	1.383
G4 (IN_ NN*)	5	24.504	3.531
G8 (VB*_ IN)			
VB*_ on	17	47.257	2.839
VB*_in	23	27.116	-1.254
VB*_ of	1	21.126	-3.762
VB*_by	10	17.147	2.170

Collocational pattern	Frequency	Likelihood	Effect
VB*_with	9	14.820	2.121

Table (4) provides an analysis of collocational patterns in Native Linguistics sub-corpus:

1. Focusing on various combinations of nouns, verbs, adjectives, and prepositions. The most frequent pattern is G1 (NN*_IN), where nouns are followed by prepositions. Specifically, nouns followed by "of" (NN*_of) are the most common, with a high likelihood and effect, indicating a strong tendency for this collocation to occur. Nouns followed by "in" (NN*_in), "for" (NN*_for), "by" (NN*_by), and "with" (NN*_with) also show a positive effect, indicating a preference for these collocations

2. The G3 (NN_ that) pattern shows a moderate likelihood and effect, suggesting a tendency for nouns to be followed by "that". The G4 (IN_ NN*) pattern reveals a high likelihood and effect, indicating a strong tendency for prepositions to be followed by nouns.

3. The G8 (VB*_ IN) pattern shows a mixed effect of verbs followed by prepositions. While verbs followed by "on" (VB*_ on) and "by" (VB*_by) have a high effect, verbs followed by "in" (VB*_in) and "of" (VB*_ of) have a negative effect, indicating a weaker tendency for these collocations to occur. Verbs followed by "with" (VB*_with) show a moderate effect, suggesting a preference for this collocation. Native speakers in linguistics sub-corpus prioritize accuracy and precision in their grammar, with nouns dominating and prepositions like "of" being prevalent. These prepositions express unique meanings and are essential for precise language communication. Verbs after prepositions vary, with "in" being

less prevalent and "on" and "by" common for significant language actions. Prepositional adjectives are sparsely used.

Table (5). Examples on Grammatical collocation patterns in Native- Linguistics sub-corpus

Label	Syntactic Pattern	Examples
G1 (NN*_IN)		
G1	NN*_of	study of syntax
	NN*_in	differences in dialects
	NN*_for	evidence for this claim
	NN*_by	influenced by culture
	NN*_with	compared with English
G3	(NN_ that)	evidence that
G4	(IN_ NN*)	in linguistics
G8 (VB*_ IN)		
G8	VB*_ on	focus on morphology
	VB*_in	engage in research
	VB*_ of	aware of the problem
	VB*_by	influenced by other
	VB*_with	compare with other languages

Table (6). Grammatical collocation patterns in Native- Translation sub-corpus

Collocational pattern	Frequency	Likelihood	Effect
G1 (NN*_ IN)			
NN*_of	160	526.608	2.489
NN*_for	21	28.558	1.695
G4 (IN_ NN*)	117	355.469	2.778
G5 (JJ_ IN)	22	22.629	-1.545
G8 (VB*_ IN)			
VB*_in	25	14.423	-0.916
VB*_by	9	17.642	2.365
VB*_with	10	28.377	2.912

The table above provides an analysis of collocational patterns in Native Translation sub-corpus, showing that:

1. The most frequent pattern is G1 (NN*_ IN), where nouns are followed by prepositions. Specifically, nouns followed by "of" (NN*_of) are the most common, with a high likelihood and effect, indicating a strong tendency for this collocation to occur. Nouns followed by "for" (NN*_for) also show a positive effect, indicating a preference for this collocation.

2. The G4 (IN_ NN*) pattern reveals a high likelihood and effect, indicating a strong tendency for prepositions to be followed by nouns. The G5 (JJ_ IN) pattern, however, shows a negative effect, suggesting that adjectives followed by prepositions are less common or have a weaker impact on the meaning or structure of the text.

3. The G8 (VB*_ IN) pattern shows a mixed effect of verbs followed by prepositions. While verbs followed by "by" (VB*_by) and "with" (VB*_with) have a high effect, verbs followed by "in" (VB*_in) have a negative effect, indicating a weaker tendency for this collocation to occur.

The table reveals native speakers' grammar usage in translation sub-corpus. Prepositions, particularly "of," are common, indicating precise definitions are essential for accurate translations. "For" after nouns indicates the target language's goal. Prepositions followed by adjectives prioritize the intended audience's needs. Prepositions and verbs have different effects, with "in" less common but crucial for defining translation tools.

Table (7). Examples on Grammatical collocation patterns in Native- Translation sub-corpus

Label	Syntactic Pattern	Examples
G1 (NN*_ IN)		
G1	NN*_of	success of any translation
	NN*_for	reasons for why audiovisual translation is so complex
G4	(IN_ NN*)	in the voice of the translator
G5	(JJ_ IN)	significant in how faithfully that narrative is presented
G8 (VB*_ IN)		
G8	VB*_in	displayed in more than one way
	VB*_by	controlled by syntactic categories
	VB*_with	complete with knowledge

4.2.2. Non- Native Corpora

Table (8) Grammatical Collocational Pattern Non- Native Literature Sub-Corpora

Collocational pattern	Frequency	Likelihood	Effect
G1 (NN*_ IN)			
NN*_of	125	652.117	3.604
G2 (NN_ to infinitive)	10	24.950	2.730
G5 (JJ_ IN)	9	31.642	3.294
G8 (VB*_ IN)			
VB*_ on	10	24.950	2.730
VB*_by	9	31.642	3.294
VB*_with	10	29.624	3.006

Table (8), represents the frequency, likelihood and effect analysis of collocation patterns in non- native sub-corpus. It shows that:

1. G1 (NN*_of): this pattern reveals a significant association between nouns and the preposition "of" in the corpus. With a high frequency of occurrence and a likelihood metric of 652.117, it indicates a prevalent syntactic structure within the non-native literature. The substantial effect value (3.604) underscores the deliberate linguistic choice or stylistic preference for this collocational pattern. It

suggests that the use of "of" following nouns is a common syntactic feature in this corpus, potentially contributing to thematic cohesion or descriptive richness.

2. G2 (NN_ to infinitive): G2, representing nouns followed by "to" and an infinitive, occurs less frequently compared to G1 but still exhibits a notable likelihood (24.950). Despite its lower frequency, the significant effect value (2.730) suggests a meaningful association between nouns and infinitive constructions in the corpus. This pattern may serve specific narrative or discourse functions, such as indicating actions, intentions, or concepts. Its intentional usage contributes to the structural organization and semantic clarity of the text.

3. G5 (JJ_ IN): which signifies adjectives followed by prepositions, a pattern observed with a relatively low frequency but notable likelihood (31.642). Despite its infrequent occurrence, the significant effect value (3.294) suggests a deliberate linguistic choice or stylistic feature in the non-native literature corpus. The association between adjectives and prepositions enriches the text's descriptive depth and expressive range, contributing to mood setting, character portrayal, or scene depiction.

4. G8 (VB*_ IN): which represents verbs followed by prepositions, observed with moderate frequency and a substantial likelihood (24.950). Its significant effect value (2.730) indicates purposeful language use, possibly for conveying actions, relationships, or spatial/temporal concepts within the narrative. The intentional deployment of verbs and prepositions enhances narrative coherence, semantic precision, and thematic development in the non-native literature corpus.

The non-native literature sub-corpus values precision and in-depth explanations, with a preference for words followed by "of" (G1). Adjectives with prepositions and infinitive verb phrases after nouns (G2) are used to incorporate subtleties and

descriptive features, while verbs followed by prepositions express relationships and activities in the story.

Table (9) Examples on Grammatical Collocational Pattern Non- Native Literature Sub-Corpora

Label	Syntactic Pattern	Examples
		(NN*_ IN)
G1	NN*_of	study of works
G2	(NN_ to infinitive)	study to examine
G5	(JJ_ IN)	study of works
		(VB*_ IN)
G8	VB*_ on	weaves narrative on characters
	VB*_by	includes analysis by applying approach
	VB*_with	mixes discrimination matters with comedy

Table (10). Grammatical collocational pattern Non-Native linguistics sub-corpora

Collocational pattern	Frequency	likelihood	Effect
G1 (NN*_ IN)			
NN*_of	66	229.810	2.574
G5 (JJ_ IN)	1	20.628	-3.703
G8 (VB*_ IN)			
VB*_by	7	31.466	3.634

Table (10) above represents the frequencies of use of grammatical collocation patterns in non- native Linguistics sub-corpus, that illustrates the following:

1. G1 (NN*_ IN) - NN*_of: The pattern "NN*_of" indicates a strong association between nouns and the preposition "of" within this corpus. It occurs frequently (66 times), suggesting a common syntactic structure. The high likelihood (229.810) indicates that this association is statistically significant, implying intentional language use. Moreover, the positive effect value (2.574) suggests a

meaningful and deliberate association between nouns and "of" in this corpus, possibly contributing to descriptive richness or thematic coherence.

2. G5 (JJ_ IN): which represents adjectives followed by prepositions, occurring only once in the corpus. Despite its rarity, the relatively high likelihood (20.628) suggests a notable association between adjectives and prepositions. The negative effect value (-3.703) is unexpected and indicates a significant but inverse association between adjectives and prepositions. This deviation from typical usage may reflect a unique stylistic choice or linguistic anomaly within the corpus.

3. G8 (VB*_ IN) - VB*_by: that denotes verbs followed by the preposition "by," observed seven times in the corpus. The likelihood (31.466) indicates a notable association between verbs and "by," suggesting intentional language use. The positive effect value (3.634) signifies a significant and deliberate association between verbs and "by," possibly indicating agency, means, or instrumentality in the corpus.

The non-native linguistics sub-corpus focuses on distilling essential ideas rather than providing detailed descriptions. They prefer precise definitions over rich descriptions, with noun sentences ending in "of" and adjective phrases with prepositions being uncommon. Verbs followed by "by" are important for language analysis, making their writing focused and succinct.

Table (11). Examples on Grammatical collocational pattern Non- Native linguistics sub-corpora

Label	Syntactic Pattern	Examples
	(NN*_ IN)	
G1	NN*_of	use of Neo-Aramaic
G5	(JJ_ IN)	positive attitudes in their ethnic language
	(VB*_ IN)	
G8	VB*_by	supported by factors

Table (12). Grammatical collocation patterns in Non-Native- Translation sub-corpus

Collocational pattern	Frequency	likelihood	Effect
G1 (NN*_ IN)			
NN*_of	87	219.242	2.201
G8 (VB*_ IN)			
VB*_ on	6	18.232	2.987
VB*_by	9	33.644	3.309

Results from non- native Translation sub-corpus show that:

1. G1 (NN*_ IN) - NN*_of: that signifies nouns followed by the preposition "of" in the corpus. It appears frequently (87 times), indicating a common syntactic structure. The relatively high likelihood (219.242) suggests a statistically significant association between nouns and "of," indicating intentional language use. The positive effect value (2.201) indicates a meaningful association between nouns and "of," possibly contributing to descriptive richness or thematic cohesion within this corpus.

2. G8 (VB*_ IN) - VB*_on and VB*_by: These patterns denote verbs followed by prepositions "on" and "by" respectively. They occur with moderate frequency (6 and 9 times), indicating intentional language use. The likelihood values (18.232 for VB*_on and 33.644 for VB*_by) suggest notable associations between verbs and their respective prepositions. The effect values (2.987 for VB*_on and 3.309 for VB*_by) indicate significant and deliberate associations between verbs and the prepositions "on" and "by," possibly conveying spatial relationships, means, or instrumentality within this corpus.

The non-native translation sub-corpus balances capturing the subtleties of the original text with being understandable. It emphasizes precise definitions with noun

phrases like "of" and uses prepositions like "by" to express acts, instruments, or reasons for certain translations. This writing style values precise translation while promoting straightforward communication.

Table (13). Examples on Grammatical collocation patterns in Non-Native- Translation sub-corpus

Label	Syntactic Pattern	Examples
(NN*_ IN)		
G1	NN*_of	translation of English novels
G8 (VB*_ IN)		
G8	VB*_ on	sheds light on the strategies
	VB*_by	manipulated by translators

4.3. Results related to Question Two "What are the similarities and differences between the use of collocations by native and non-native student writers across genres (linguistics, literature, and translation studies)?"

To answer this question, the researcher established a comparison between native and non-native sub-corpora of each filed: native literature vs. non- native literature, native linguistics vs. non-native linguistics sub-corpora, and native translation vs. non-native translation sub-corpora.
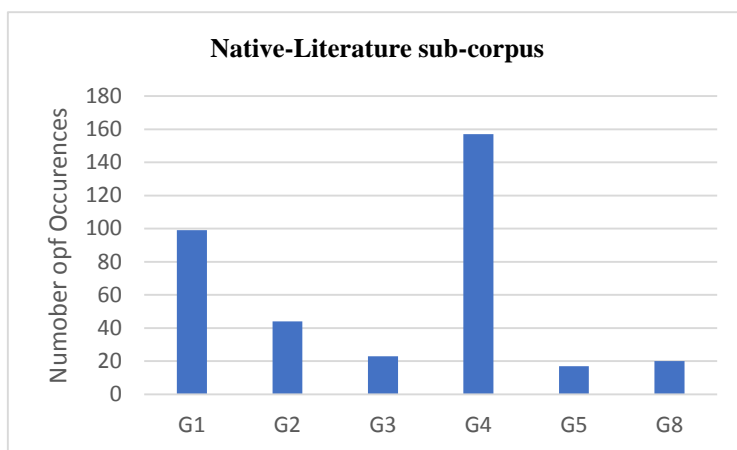
Figure (1). Count Frequency of Grammatical Collocation Patterns in Native literature sub- corpus
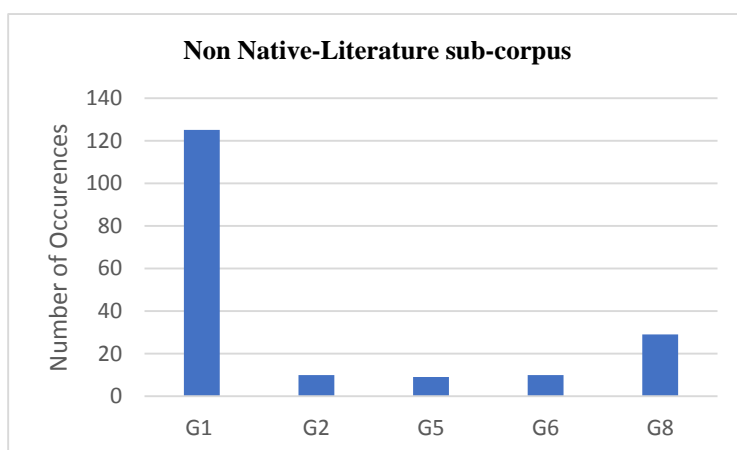
Figure (2). Count Frequency of Grammatical Collocation Patterns in non- native literature sub- corpus

Figure 1 and 2 are demonstrate the similarities and differences in use of grammatical collocation patterns between Native and non- native literature sub-corpora. When comparing the native literature sub-corpus with the non-native literature sub-corpus, both exhibit similarities and differences in their collocational patterns. Similarities are evident in the presence of common patterns (G1, G2, G5, G8), indicating shared linguistic structures across both types of literature. However, differences emerge in several aspects. Firstly, the frequency distribution of collocational patterns varies between the two sub-corpora, with differing prevalence and usage observed.

Moreover, the native literature sub-corpus includes unique patterns (G3 and G4) not found in the non-native literature sub-corpus, suggesting distinct linguistic features or stylistic preferences. Furthermore, for corresponding patterns present in both sub-corpora, such as G1 and G2, significant differences in frequency counts highlight varying degrees of prevalence or usage between native and non-native literature.

Lastly, the total number of patterns differs between the two sub-corpora, with the non-native literature sub-corpus containing an additional pattern (G6), indicative of potential linguistic variations or stylistic preferences unique to non-native literary texts. These similarities and differences contribute to a deeper understanding of the linguistic

Native Linguistics sub-corpus
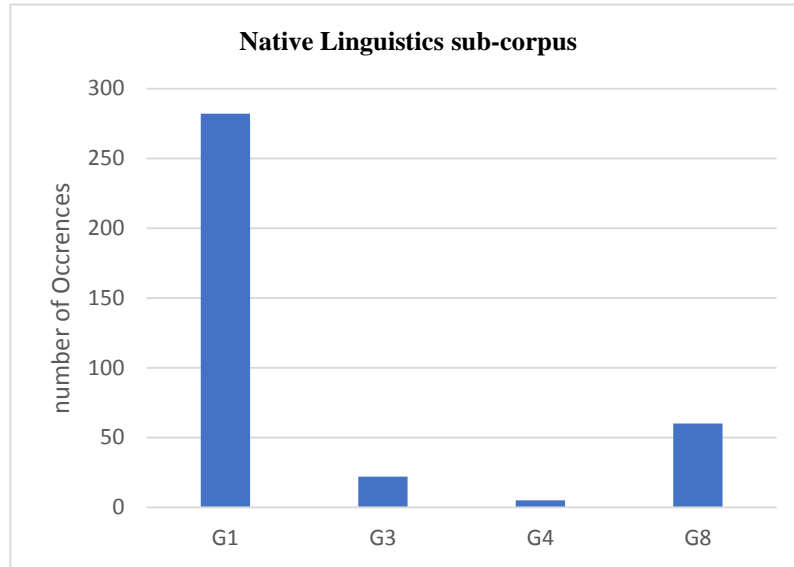
Figure (3). Count Frequency of Grammatical Collocation Patterns in Native Linguistics sub-corpus

Non-native Linguistics sub-corpus

Figure (4). Count Frequency of Grammatical Collocation Patterns in non- native Linguistics sub-corpus

When comparing the grammatical collocation patterns utilized in the native linguistics sub-corpus with those in the non-native linguistics sub-corpus, both similarities and differences become apparent. Across both sub-corpora, commonalities

are observed in the presence of shared collocational patterns: G1 and G8, indicating overlapping linguistic structures between native and non-native linguistic texts.

However, notable differences emerge in various aspects. Firstly, there is a stark contrast in frequency distribution, with the native linguistics sub-corpus displaying a wider range of frequencies (from 5 to 282 occurrences per pattern) compared to the non-native linguistics sub-corpus, where frequencies are generally lower (ranging from 1 to 66 occurrences per pattern).

Furthermore, the native linguistics sub-corpus includes patterns, specifically G3 and G4, which are absent in the non-native linguistics sub-corpus, suggesting distinct linguistic phenomena or structures specific to native linguistic texts. Moreover, for patterns present in both sub-corpora, such as G1 and G8, significant differences in frequency counts highlight varying prevalence or usage between native and non-native linguistic texts. These disparities in frequency distribution, presence of some patterns, and frequency disparity for corresponding patterns underscore the nuanced differences in linguistic characteristics between native and non-native linguistic texts.
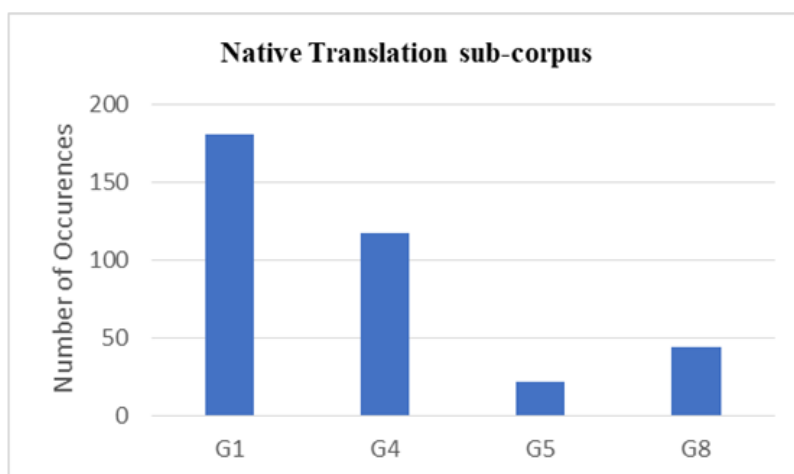
Figure (5). Count Frequency of Grammatical Collocations in Native Translation sub-corpus

Non- native Translation sub-corpus
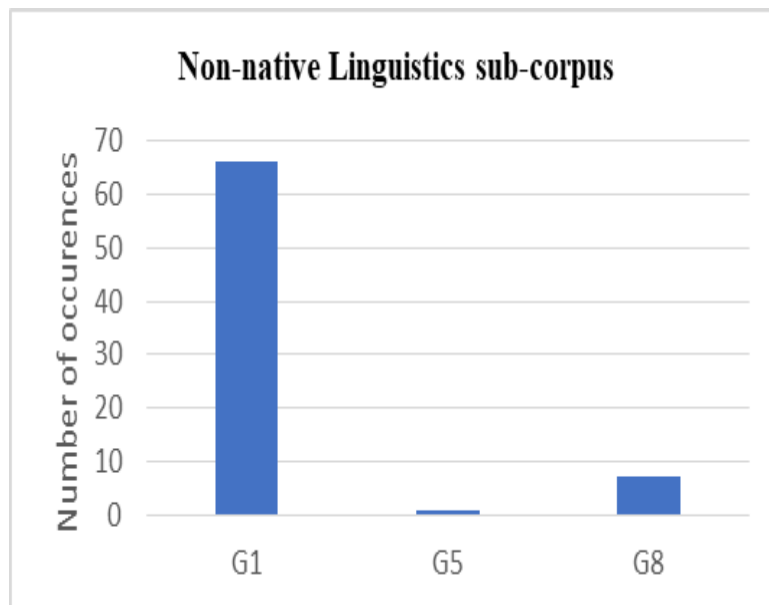
Number of Occurences

100
80
60
40
20
0

G1   G8
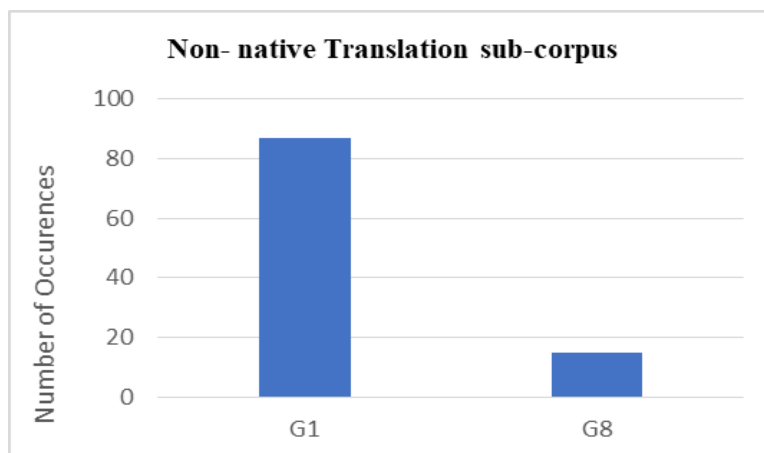
Figure (6). Count Frequency of Grammatical Collocations in non- native Translation sub-corpus

When comparing the grammatical collocation patterns present in the native translation sub-corpus with those in the non-native translation sub-corpus, both similarities and differences are discernible. Shared collocational patterns G1 and G8 are evident in both sub-corpora, suggesting common linguistic structures across native and non-native translation texts. However, notable differences arise in various aspects. While the native translation sub-corpus demonstrates a wide range of frequencies for patterns, ranging from 22 to 181 occurrences per pattern, the non-native translation sub-corpus exhibits lower frequencies, typically ranging from 15 to 87 occurrences per pattern.

Moreover, patterns as G4 and G5 are exclusively found in the native translation sub-corpus, indicating linguistic structures or phenomena specific to native translation texts. Further, for shared patterns between the two sub-corpora, such as G1 and G8, significant differences in frequency counts underscore varying prevalence or usage in native versus non-native translation texts. These discrepancies reflect nuanced differences in linguistic characteristics, stylistic conventions, and thematic elements

inherent to each type of text, enriching our understanding of translation processes and linguistic dynamics across native and non-native contexts.

In summary, examining native and non-native language use in literature, linguistics, and translation sub-corpora reveals both similarities and differences. While some patterns are shared, like G1 and G8, there are notable variations in how often they appear and in the presence of other patterns. These differences highlight the diverse ways language is used and understood in native versus non-native contexts. Understanding such information helps us grasp the complexities of language, translation, and cultural expression across different settings, enriching our appreciation of linguistic diversity and its role in shaping communication and understanding. Yet, the level of competency also might play a significant role when using collocations especially that they convey non- compositional meaning, which might be challenging for second language learners.

CHAPTER FIVE

Conclusions, and Recommendations

5.1. Conclusion

This section offers an analysis and interpretation of the collocational patterns detected in both native and non-native sub-corpora within the fields: literature, linguistics, and translation. By conducting a corpus driven approach, the investigation of these patterns highlights the variations and similarities in the utilization of grammatical collocation patterns, thus enhancing the comprehension of linguistic dynamics across various contexts. Comprehension of linguistic dynamics across various contexts.

5.1.1. Common Collocational Patterns Across Genres by native and non- native users

The examination unveiled the existence of mutual collocational patterns, specifically G1 (noun-preposition) and G8 (verb-preposition), within native and non-native sub-corpora. This finding implies the presence of particular linguistic formations and practices that go beyond genre distinctions, demonstrating universal trends in linguistic utilization and representation. The frequency of these shared patterns highlights their importance in influencing communication across varying linguistic environments.

5.1.2. Variations in Frequency Distribution

Significant differences in the frequency distribution of collocational patterns were observed between native and non-native sub-corpora. Such variations may reflect varying usage of collocations across different linguistic contexts, potentially influenced

by factors such as language proficiency, cultural background, and communicative norms. While some patterns may be more dominant in native corpora, others may be more prominent in non-native texts, highlighting the diverse ways in which language is used and understood. This outcome is compatible with Liu's and Jiang's (2018) analysis, who conducted a coprus-based examination which uncovered some differences in collocational patterns between native and non-native English writing. According to them, native speakers displayed a broader array of collocational options in contrast to non-native speakers, indicating variation in language proficiency and communicative effectiveness among the two groups. This result corresponds with the concept that differences in language proficiency impact the frequency distribution of collocational patterns, with native speakers likely showing a more extensive range of linguistic expressions.

In addition to that, Paquot and Granger (2012) investigated verb-noun collocations in native and non-native English academic writing and found distinct patterns of collocational usage between the two groups. While certain collocations were more prevalent in native writing, others were more prominent in non-native writing. These variations were attributed to differences in linguistic background, cultural influence, and language learning experience. This study highlights the role of cultural and linguistic factors in shaping collocational usage, contributing to the observed differences in frequency distribution.

Similarly, Zareva (2017) analyzed how noun phrases are used in English academic writing by both native and non-native speakers. She noted variations in collocation frequency distribution, finding that native writers had more diverse choices than non-native writers. Variations in language proficiency and exposure to academic discourse

conventions were connected to these differences. Zareva's research supports the idea that variations in language skills and communication norms impact the frequency and application of collocational patterns in various linguistic environments.

5.1.3. Unique Patterns in Native Sub-Corpora

It was evident that native sub-corpora exhibited unique collocational patterns, such as G3 and G4, which were not found in non-native sub-corpora. These distinctive patterns suggest the presence of distinct linguistic features or stylistic preferences specific to native texts, contributing to the richness and diversity of native language use. The identification of these unique patterns underscores the importance of considering native language conventions in linguistic analysis and interpretation. Such analysis agrees with Lee's and Swales'(2006) study that examined academic writing by native and non-native English speakers and identified unique collocational patterns in native texts. They found that native speakers employed a wider range of discourse markers and idiomatic expressions compared to non-native speakers, suggesting the presence of distinct linguistic features specific to native language use.

On the other hand, the results of the current analysis contrast with other studies such as Granger (1998), who examined the use of collocations in native and non-native English learner writing and found that while non-native speakers exhibited a narrower range of collocational patterns compared to native speakers, they sometimes demonstrated creativity in their use of language. Non-native texts occasionally featured innovative collocational combinations that were not.

5.1.4. Salient Frequencies

As the analysis demonstrated, the differences in the total number of collocational patterns between native and non-native sub-corpora were also evident. Non-native sub-corpora often contained additional patterns, such as G6, indicating potential linguistic variations or stylistic preferences unique to non-native texts. In addition to that, the analysis showed that within the linguistics genre, a clear contrast in frequency distribution was noted between native and non-native sub-corpora. Where the native corpus exhibited a wider range of frequencies compared to non-native texts, reflecting varying degrees of linguistic complexity and sophistication. These differences may be attributed to differences in language proficiency and communicative competence, as well as variations in academic writing conventions across different linguistic contexts.

5.1.5. Exclusive Presence of Patterns in Native Translation Sub-Corpus

It was noticed that certain collocational patterns were exclusively found in the native translation sub-corpora, indicating linguistic structures specific to native translation texts, which are G4 and G5. The presence of these exclusive patterns highlights the unique linguistic characteristics and stylistic conventions inherent in native translation practices. These findings emphasize the importance of considering translation as a complex linguistic and cultural phenomenon shaped by native language conventions and translation norms.

In conclusion, the examination of collocational patterns in native and non-native sub-corpora across various genres brings to light several significant discoveries. First, there are shared collocational patterns evident in both native and non-native texts, indicating universal linguistic tendencies. Subsequently, noteworthy disparities in the distribution frequency of collocational patterns point towards varying prevalence and

usage shaped by elements like language proficiency and cultural background. These outcomes are consistent with prior research and underscore the necessity of taking into account linguistic differences during analysis. Further, distinctive collocational patterns unique to native texts highlight specific linguistic attributes inherent in native language usage. Moreover, discrepancies in the overall quantity of patterns between native and non-native texts demonstrate differences in linguistic intricacy. Finally, specific collocational patterns exclusive to native translation sub-corpora shows distinct linguistic features essential to translation methodologies. Thus, these findings enrich our comprehension of language diversity and underscore the significance of considering linguistic and cultural elements in linguistic examination.

5.2. Recommendations

This study recommends the following:

1. To conduct further analysis on the use of collocations in context, focusing on different genres, which strengthen ESP teaching and learning.

2. To conduct more studies on second language learners' errors when using collocations whether they are lexical or grammatical in the spoken and written contexts.

3. To conduct studies on the Jordanian context detecting if Jordanian English second language learners tend to invent collocation patterns due to the influence of their first language and studying the spread of such phenomenon among Jordanian students.

4. To conduct a corpus analysis of collocational patterns in spoken and written English   by public and private university students in Jordan.

References

Ackermann, Kirsten & Chen, Yu-Hua. (2013). Developing the Academic Collocation List (ACL) – A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes*. 12. 235–247. 10.1016/j.jeap.2013.08.002.

Anthony, L. (2019). AntConc [Computer software]. https://programminghistorian.org/en/lessons/corpus-analysis-with-antconc

Benson, M., Benson, E., & Ilson, R. F. (1986). Lexicographic description of English (Vol. 14). John Benjamins Publishing.

Chang, Y. (2018). Features of Lexical Collocations in L2 Writing: A Case of Korean Adult Learners of English. English Teaching, 73(2), 3-36.

Demir, C. (2017). Lexical collocations in English: A comparative study of native and non-native scholars of English. *Journal of Language and Linguistic Studies*, 13(1), 75-87.

Du, X & Afzaal, M & Alfadda, H. (2022). *Collocation Use in EFL Learners' Writing Across Multiple Language Proficiencies: A Corpus-Driven Study. Frontiers in Psychology.* 13. 10.3389/fpsyg.2022.752134.

Dukali, A . (2016). Difficulties Encountered by Libyan University Students of English as a Foreign Language in the Use of Lexical Collocations. Doctoral thesis, University of Huddersfield. This version is available at http://eprints.hud.ac.uk/id/eprint/34435/

Duong, Duc & Nguyen, Nha. (2021). Using Collocations to Enhance Academic Writing: A Survey Study at Van Lang University. 10.2991/assehr.k.210226.035.

Durrant, Philip & Schmitt, Norbert. (2009). To what extent do native and non-native writers make use of collocations?. *Iral-international Review of Applied Linguistics in Language Teaching* - IRAL-INT REV APPL LINGUIST. 47. 157-177. 10.1515/iral.2009.007.

Farghal, M., & Al-Hamly, M. (2007). Lexical collocations in EFL writing. *Journal of Asia TEFL*, 4(2).

Farooqui, A. (2016). A corpus-based study of academic-collocation use and patterns in postgraduate Computer Science students' writing. PhD thesis, University of Essex.

Firth, J. (1957). A Synopsis of Linguistic Theory, 1930-55. In Studies in Linguistic Analysis (pp. 1-31). Special Volume of the Philological Society. Oxford: Blackwell. [Reprinted as Firth (1968)]

Froehlich, H. (2015). Corpus Analysis with Antconc. The Programming Historian. 10.46430/phen0043.

Futagi, Y., Deane, P., Chodorow, M., & Tetreault, J. (2008). A computational approach to detecting collocation errors in the writing of non-native speakers of English. Computer Assisted Language Learning, 21(4), 353-367.

Granger, S. (1998). *Prefabricated Patterns in Advanced EFL Writing: Collocations and Formulae.* 10.1093/oso/9780198294252.003.007.

Gries, S. (2009). What is corpus linguistics?. Language and linguistics compass, 3(5), 1225-1241.

Hardie, A. (2015). *Corpus linguistics. In The Routledge handbook of linguistics* (pp. 502-515). Routledge.

Holtz, M. (2011). *Lexico-grammatical properties of abstracts and research articles. A corpus-based study of scientific discourse from multiple disciplines*. (full text in English:https://tuprints.ulb.tu-darmstadt.de/2638/1/PhD-Thesis-Monica-Holtz.pdf).

Lee, David & Swales, John. (2006). A corpus-based EAP COURSE for NNS doctoral students: moving from available specialized corpora to self-compiled corpora. English for Specific Purposes. 25. 56-75. 10.1016/j.esp.2005.02.010.

Lei, L & Liu, D. (2018). *The academic English collocation list: A corpus-driven study. International Journal of Corpus Linguistics,* 23, 0216-243. 10.1075/ijcl.16135.lei.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: A practical guide. Routledge.*

Mirsalari, S. (2019). An Investigation into the Developmental Patterns of Lexical Collocation among Iranian EFL Learners. *Applied Linguistics Research Journal*. 3. 10.14744/alrj.2019.02486.

Moehkardi, Rio & Moehkardi, Diah. (2002). grammatical and lexical English collocations: some possible problems to Indonesian learners of English. *Jurnal Humaniora*, 14, 53-62   http://dx.doi.org/10.22146/jh.745

Mousavi, S. M., & Darani, L. H. (2018). *Effect of collocations on Iranian EFL learners' writing: Attitude in focus. Global Journal of Foreign Language Teaching,* 8(4), 131-145. http://dx.doi.org/10.18844/gjflt.v8i4.3568

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

O'Flynn, J. (2018, September). *Developing an Academic Collocation List for Arts and Humanities*[Master's dissertation, University of Warwick] .https://www.teachingenglish.org.uk/sites/teacheng/files/James%20O%27Flynn_ University%20of%20Warwick_Dissertation.pdf

Orol, Ana & Ramos, Margarita. (2013). A Comparative Study of Collocations in a Native Corpus and a Learner Corpus of Spanish. Procedia - Social and Behavioral Sciences. 95. 563-570. 10.1016/j.sbspro.2013.10.683.

Paquot, Magali & Granger, Sylviane. (2012). Formulaic Language in Learner Corpora. *Annual Review of Applied Linguistics*. 32. 130-149. 10.1017/S0267190512000098.

Sari, Beaty & Gulö, Ingatan. (2019). Observing Grammatical Collocation in Students' Writings. TEKNOSASTIK. 17. 25. 10.33365/ts.v17i2.325.

Seretan, V. (2005, July). Induction of syntactic collocation patterns from generic syntactic relations. *In International Joint Conference on Artificial Intelligence*, 19, 1698. *Lawrence Erlbaum Associates*. LTD.

Sinclair, J. (Ed.) (1991). *Corpus, Concordance, Collocation*. Oxford, UK: Oxford University Press.

Stubbs, M. (2008). *Text and corpus analysis: Computer-assisted studies of language and culture. Routledge*

Suraprajit, P. (2022). *English words and collocations found on logistics magazines: A corpus-based study*. Advances in Language and Literary Studies, 13(1), 41-48. ., 10.7575/aiac.alls.v.13n.1.p.41.

Thongvitit, S., & Thumawongsa, N. (2017). *A corpus-based study of English collocations found in the abstracts of research articles written by Thai EFL writers.* International Journal of Social Science and Humanity, 7(12), 751-755.

Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work. Computational Linguistics*, 28, 583-583. 10.1162/coli.2002.28.4.583a.

Wongwiwat, T. (2016). *Move analysis and lexical bundle analysis of conference abstracts: A case study of Thailand TESOL international conferences.* Unpublished doctoral dissertation, Thammasat University.

Wu, Yong & Abd Halim, Hasliza. (2022). A Systematic Review of Studies on L2 Collocational Knowledge among Non-native English Learners. 22. 755-772. 10.37178/ca-c.21.5.061.

Zareva, Alla. (2020). *Collocations in Student Academic Presentations*. 10.1007/978-3-030-37980-3_4. Retrieved from

Zou, B., & Thomas, M. (Eds.). (2018). *Handbook of research on integrating technology into contemporary language learning and teaching.* IGI Global.

Appendices

Appendix (1)

AntConc Picture Example

G1 patterns in native Linguistics sub-corpus:

	File	Left Context	Hit	Right Context
1	LINGUISTI...	NN individual_NN ,_, contrasting_VBG identities_NNS in_IN Japanese_	NNP	and_CC English_NNP in_IN a_DT Japanese_
2	LINGUISTI...	DT Jewish_JJ community_NN of_IN Dayton_NNP ,_, Ohio_	NNP ,_,	and_CC how_WRB features_NNS of_IN Yiddish_
3	LINGUISTI...	IN North_NNP Dakota_NNP ,_, South_NNP Dakota_NNP ,_, Iowa_	NNP ,_,	and_CC Illinois_NNP ._. Initially_RB led_VBN by_
4	LINGUISTI...	VBP this_DT approach_NN using_VBG a_DT Word_	NNP	and_CC Paradigm_NNP model_NN incorporating_VBG multidimensional_
5	LINGUISTI...	IN a_DT phrase_NN ,_, in_IN Jewish_JJ English_	NNP ,_,	and_CC particularly_RB in_IN the_DT variety_
6	LINGUISTI...	JJ patterns_NNS of_IN the_DT Modern_NNP Greek_	NNP	and_CC Russian_JJ gaps_NNS resemble_VBP those_
7	LINGUISTI...	JJ correlation_NN between_IN presence_NN of_IN jargon_	NNP	and_CC self_NN -_HYPH assessment_NN of_IN
8	LINGUISTI...	DT university_NN in_IN the_DT United_NNP States_	NNP	and_CC sixty_CD -_HYPH two_CD undergraduate_JJ
9	LINGUISTI...	CC methods_NNS of_IN Critical_NNP Discourse_NNP Analysis_	NNP	and_CC Systemic_NNP Functional_NNP Linguistics_NNP ,_, specifically_
10	LINGUISTI...	NN and_CC function_NN in_IN Jewish_JJ English_	NNP ,_,	and_CC that_IN Yiddish_NNP was_VBD the_
11	LINGUISTI...	RB found_VBN between_IN presence_NN of_IN jargon_	NNP	and_CC the_DT ability_NN to_TO correctly_
12	LINGUISTI...	NN analyses_NNS showed_VBD that_IN learners_NNS '_" L1_	NNP	and_CC their_PRP$ learning_NN environment_NN had_
13	LINGUISTI...	IN society_NN is_VBZ (_-LRB- Joseph_NNP ,_, 2010_CD ;_: Omoniyi_	NNP ,_, 2006_	CD)_-RRB- ._. However_RB ,_, as_IN human_JJ beings_
14	LINGUISTI...	WP they_PRP are_VBP (_-LRB- Llamas_NNP &_CC Watt_	NNP ,_, 2010_	CD)_-RRB- ._. It_PRP is_VBZ through_IN language_

G8 patterns in native Linguistics sub-corpus:

#	File	Left Context	Hit	Right Context
1	LINGUISTI...	VBN evidences_NNS of_IN things_NNS which_WDT are_	VBP	not_RB easily_RB learned_VBN from_IN distributional_
2	LINGUISTI...	NNS where_WRB high_JJ frequency_NN markers_NNS are_	VBP	not_RB enough_JJ for_IN the_DT formation_
3	LINGUISTI...	VBD to_TO show_VB that_IN humans_NNS are_	VBP	not_RB sensitive_JJ to_IN verbal_JJ subcategorization_
4	LINGUISTI...	JJ systems_NNS ,_, in_IN ways_NNS that_WDT are_	VBP	not_RB typically_RB acknowledged_VBN even_RB by_
5	LINGUISTI...	CC character_NN role_NN ._. Chinese_JJ writers_NNS rely_	VBP	on_IN figurative_JJ languages_NNS to_TO embellish_
6	LINGUISTI...	NN is_VBZ often_RB overlooked_VBN ._. Researchers_NNS focus_	VBP	on_IN interviews_NNS with_IN the_DT parents_
7	LINGUISTI...	NN ._. To_TO do_VB so_RB ,_, I_PRP draw_	VBP	on_IN tools_NNS and_CC methods_NNS of_
8	LINGUISTI...	IN the_DT inflectional_JJ paradigm_NN ._. I_PRP formalize_	VBP	this_DT approach_NN using_VBG a_DT Word_
9	LINGUISTI...	TO investigate_VB the_DT factors_NNS that_WDT influence_	VBP	this_DT process_NN ._. This_DT thesis_NN reports_
10	LINGUISTI...	NNS to_TO pattern_VB differently_RB ._. I_PRP interpret_	VBP	this_DT to_TO mean_VB that_IN the_
11	LINGUISTI...	NN forms_NNS :_: words_NNS ,_, characters_NNS ,_, images_NNS ,_, are_	VBP	used_VBN for_IN inducing_VBG clusters_NNS for_
12	LINGUISTI...	WRB features_NNS of_IN Yiddish_JJ intonation_NN are_	VBP	used_VBN in_IN Jewish_JJ English_NNP ._. 20_CD
13	LINGUISTI...	VBP those_DT which_WDT previous_JJ researchers_NNS have_	VBP	used_VBN to_TO posit_VB that_IN gaps_
14	LINGUISTI...	JJ and_CC progressive_JJ tenses_NNS ,_, students_NNS are_	VBP	given_VBN either_CC Traditional_NNP Grammar_NNP instruction_

G1 patterns in Non-Native Linguistics sub-corpus

#	File	Left Context	Hit	Right Context
1	linguistis ...	NNS involved_VBN in_IN using_VBG Arabic_NNP ,_, English_	NNP	and_CC French_NNP ,_, and_CC Arabs_NNPS '_POS
2	linguistis ...	NN ._. In_IN addition_NN ,_, they_PRP use_VBP English_	NNP	and_CC French_NNP in_IN Governmental_JJ offices_
3	linguistis ...	IN Quebec_NNP have_VBP towards_IN Arabic_NNP ,_, English_	NNP	and_CC French_NNP ._. As_IN for_IN the_
4	linguistis ...	NNS and_CC phrases_NNS of_IN both_DT SL_	NNP	and_CC TL_NNP in_IN the_DT same_
5	linguistis ...	IN the_DT meaning_NN of_IN the_DT SL_	NNP	and_CC TL_NNP texts_NNS are_VBP concerned_
6	linguistis ...	the_DT use_NN of_IN Neo_NNP -_HYPH Aramaic_	NNP	and_CC Arabic_NNP ._. These_DT factors_NNS were_
7	linguistis ...	IN using_VBG Arabic_NNP ,_, English_NNP and_CC French_	NNP ,_,	and_CC Arabs_NNPS '_POS attitudes_NNS towards_IN
8	linguistis ...	CD (_-LRB- 1981_CD and_CC 1986_CD)_-RRB- of_IN Government_	NNP	and_CC Binding_NNP Theory_NNP ._. The_DT researcher_
9	linguistis ...	namely_RB ,_, Al_NNP -_HYPH Wehdat_NNP Camp_NNP ,_, Sweifieh_	NNP	and_CC Jabal_NNP AlHussein_NNP ,_, and_CC the_
10	linguistis ...	Radford_NNP 's_POS (_-LRB- 1988_CD)_-RRB- and_CC Jalabneh_	NNP	and_CC Mahmoud_NNP 's_POS (_-LRB- 2005_CD)_-RRB-
11	linguistis ...	NNP Camp_NNP ,_, Sweifieh_NNP and_CC Jabal_NNP AlHussein_	NNP ,_,	and_CC the_DT number_NN of_IN shop_
12	linguistis ...	NN ;_: they_PRP were_VBD recieved_VBN in_IN English_	NNP	and_CC translated_VBN by_IN editors_NNS of_
13	linguistis ...	IN Jordan_NNP such_JJ as_IN Salt_NNP ,_, Jerash_	NNP	and_CC Zarqa_NNP ._. his_PRP$ study_NN aimed_
14	linguistis ...	WDT supported_VBD the_DT use_NN of_IN Neo_	NNP -_	HYPH Aramaic_NNP and_CC Arabic_NNP ._. These_DT

G8 patterns in Non-Native Linguistics sub-corpus

	File	Left Context	Hit	Right Context
1	linguistis ...	CD teacher_NN ._. The_DT obtained_VBN data_NNS indicate_	VBP	that_IN English_NNP is_VBZ the_DT most_
2	linguistis ...	CC bilingual_JJ signs_NNS ._. Results_NNS also_RB show_	VBP	that_IN foreign_JJ names_NNS in_IN shop_
3	linguistis ...	DT one_CD hand_NN ,_, these_DT results_NNS indicate_	VBP	that_IN the_DT spread_NN of_IN English_
4	linguistis ...	and_CC it_PRP was_VBD 0.836_CD ._. Results_NNS show_	VBP	that_IN there_EX are_VBP two_CD main_
5	linguistis ...	way_NN of_IN life_NN ?_. 2_LS ._. What_WP are_	VBP	the_DT negative_JJ influences_NNS of_IN the_
6	linguistis ...	the_DT following_VBG questions_NNS :_: 1_LS ._. What_WP are_	VBP	the_DT positive_JJ aspects_NNS of_IN the_
7	linguistis ...	IN the_DT sentences_NNS ._. The_DT findings_NNS support_	VBP	the_DT questions_NNS and_CC the_DT objectives_
8	linguistis ...	and_CC their_PRP$ languages_NNS ?_. 3_LS ._. What_WP are_	VBP	the_DT suggestions_NNS to_TO support_VB and_
9	linguistis ...	DT questions_NNS and_CC the_DT objectives_NNS have_	VBP	been_VBN proved_VBN correct_JJ ._. This_DT study_
10	linguistis ...	DT SL_NNP and_CC TL_NNP texts_NNS are_	VBP	concerned_VBN ,_, it_PRP was_VBD observed_VBN by_
11	linguistis ...	DT radio_NN ._. In_IN addition_NN ,_, they_PRP use_	VBP	English_NNP and_CC French_NNP in_IN Governmental_
12	linguistis ...	LRB- 100_CD)_-RRB- Arab_JJ respondents_NNS who_WP reside_	VBP	in_IN Quebec_NNP –_: Canada_NNP ,_, covering_VBG different_
13	linguistis ...	JJ names_NNS in_IN shop_NN signs_NNS are_	VBP	increasingly_RB used_VBN by_IN shop_NN owners_
14	linguistis ...	VBG or_CC adding_VBG categories_NNS which_WDT constitute_	VBP	separate_JJ and_CC full_JJ arguments_NNS could_

G1 patterns in Native literature sub-corpus:

	File	Left Context	Hit	Right Context
1	LITERATU...	father_NN -_HYPH daughter_NN relationships_NNS between_IN Juliet_	NNP	and_CC her_PRP$ father_NN Capulet_NNP on_
2	LITERATU...	NNP on_IN one_CD hand_NN ,_, and_CC Ophelia_	NNP	and_CC her_PRP$ father_NN Polonius_NNP on_
3	LITERATU...	JJ ._. Through_IN the_DT dialogues_NNS between_IN Ophelia_	NNP	and_CC her_PRP$ father_NN Polonius_NNP and_
4	LITERATU...	PRP$ father_NN Polonius_NNP and_CC between_IN Juliet_	NNP	and_CC her_PRP$ father_NN Capulet_NNP ,_, the_
5	LITERATU...	POS poetry_NN from_IN three_CD aspects_NNS :_: -_: Animalism_	NNP	and_CC its_PRP$ role_NN in_IN his_
6	LITERATU...	IN England_NNP after_IN the_DT Industrial_NNP Revolution_	NNP	and_CC its_PRP$ consequences_NNS ._. Being_VBG an_
7	LITERATU...	IN the_DT British_JJ occupation_NN of_IN Palestine_	NNP	and_CC its_PRP$ oppression_NN :_: he_PRP took_
8	LITERATU...	this_DT relationship_NN .In_. Hamlet_NNP and_CC Romeo_	NNP	and_CC Juliet_NNP ,_, he_PRP shows_VBZ two_
9	LITERATU...	lives_NNS of_IN their_PRP$ daughters_NNS ._. _SP Ophelia_	NNP	and_CC Juliet_NNP are_VBP characterized_VBN as_
10	LITERATU...	IN her_PRP$ social_JJ milieu_NN ._. Through_IN Ophelia_	NNP	and_CC Juliet_NNP ,_, Shakespeare_NNP shows_VBZ two_
11	LITERATU...	VBZ all_RB about_IN the_DT Temple_NNP Mount_	NNP	and_CC the_DT Wailing_VBG Wall_NNP especially_
12	LITERATU...	DT Harry_NNP Potter_NNP Series_NNP :_: Harry_NNP Potter_	NNP	and_CC the_DT Sorcerer_NNP 's_POS Stone_
13	LITERATU...	VBN as_IN vague_JJ ._. Keywords_NNS :_: Harry_NNP Potter_	NNP	and_CC the_DT Sorcerer_NNP 's_POS Stone_
14	LITERATU...	central_JJ theme_NN ,_, i.e._FW Jack_NNP Kerouac_	NNP	and_CC Hussein_NNP Mardan_NNP are_VBP both_

G8 patterns in Native literature sub-corpus:

	File	Left Context	Hit	Right Context
1	LITERATU...	NNS of_IN the_DT two_CD fathers_NNS reflect_	VBP	the_DT dominance_NN and_CC control_NN over_
2	LITERATU...	DT following_VBG three_CD questions_NNS :_: What_WP are_	VBP	the_DT elements_NNS that_WDT make_VBP Bingo_
3	LITERATU...	PRP$ family_NN in_IN general_NN ?_. What_WP are_	VBP	the_DT real_JJ drives_NNS which_WDT stand_
4	LITERATU...	NNS "_" ,_, meaning_VBG the_DT characters_NNS who_WP represent_	VBP	the_DT struggle_NN of_IN individuals_NNS in_
5	LITERATU...	PRP$ harsh_JJ reality_NN ._. The_DT plays_NNS present_	VBP	the_DT suffering_NN of_IN the_DT two_
6	LITERATU...	VBP authoritative_JJ and_CC dominating_VBG .They_PRP want_	VBP	to_TO control_VB and_CC command_VB the_
7	LITERATU...	NN ._. Nevertheless_RB ,_, the_DT three_CD poets_NNS seek_	VBP	to_TO glorify_VB Jerusalem_NNP 's_NN holiness_
8	LITERATU...	IN modern_JJ poetry_NN ._. This_DT study_NN aim_	VBP	to_TO investigate_VB modern_JJ literature_NN 's_
9	LITERATU...	NNS of_IN women_NNS which_WDT eventually_RB lead_	VBP	to_IN inflicting_VBG some_DT damaging_JJ effects_
10	LITERATU...	TO analyze_VB the_DT elements_NNS that_WDT make_	VBP	Bingo_NNP a_DT biographical_JJ fiction_NN by_
11	LITERATU...	WP are_VBP the_DT elements_NNS that_WDT make_	VBP	Bingo_NNP a_DT biographical_JJ fiction_NN ?_. How_
12	LITERATU...	IN the_DT various_JJ problems_NNS women_NNS face_	VBP	in_IN patriarchal_JJ societies_NNS ._. This_DT study_
13	LITERATU...	DT injustices_VBZ the_DT female_JJ characters_NNS suffer_	VBP	in_IN their_PRP$ man_NN -_HYPH dominated_VBN
14	LITERATU...	VBG these_DT female_JJ characters_NNS that_WDT have_	VBP	many_JJ differences_NNS and_CC gaps_NNS between_

G1 patterns in Non-Native Literature sub-corpus

	File	Left Context	Hit	Right Context
1	LITERATU...	father_NN -_HYPH daughter_NN relationships_NNS between_IN Juliet_	NNP	and_CC her_PRP$ father_NN Capulet_NNP on_
2	LITERATU...	NNP on_IN one_CD hand_NN ,_, and_CC Ophelia_	NNP	and_CC her_PRP$ father_NN Polonius_NNP on_
3	LITERATU...	JJ ._. Through_IN the_DT dialogues_NNS between_IN Ophelia_	NNP	and_CC her_PRP$ father_NN Polonius_NNP and_
4	LITERATU...	PRP$ father_NN Polonius_NNP and_CC between_IN Juliet_	NNP	and_CC her_PRP$ father_NN Capulet_NNP ,_, the_
5	LITERATU...	POS poetry_NN from_IN three_CD aspects_NNS :_: -_: Animalism_	NNP	and_CC its_PRP$ role_NN in_IN his_
6	LITERATU...	IN England_NNP after_IN the_DT Industrial_NNP Revolution_	NNP	and_CC its_PRP$ consequences_NNS ._. Being_VBG an_
7	LITERATU...	IN the_DT British_JJ occupation_NN of_IN Palestine_	NNP	and_CC its_PRP$ oppression_NN :_: he_PRP took_
8	LITERATU...	this_DT relationship_NN .In_. Hamlet_NNP and_CC Romeo_	NNP	and_CC Juliet_NNP ,_, he_PRP shows_VBZ two_
9	LITERATU...	lives_NNS of_IN their_PRP$ daughters_NNS ._. _SP Ophelia_	NNP	and_CC Juliet_NNP are_VBP characterized_VBN as_
10	LITERATU...	IN her_PRP$ social_JJ milieu_NN ._. Through_IN Ophelia_	NNP	and_CC Juliet_NNP ,_, Shakespeare_NNP shows_VBZ two_
11	LITERATU...	VBZ all_RB about_IN the_DT Temple_NNP Mount_	NNP	and_CC the_DT Wailing_VBG Wall_NNP especially_
12	LITERATU...	DT Harry_NNP Potter_NNP Series_NNP :_: Harry_NNP Potter_	NNP	and_CC the_DT Sorcerer_NNP 's_POS Stone_
13	LITERATU...	VBN as_IN vague_JJ ._. Keywords_NNS :_: Harry_NNP Potter_	NNP	and_CC the_DT Sorcerer_NNP 's_POS Stone_
14	LITERATU...	central_JJ theme_NN ,_, i.e._FW Jack_NNP Kerouac_	NNP	and_CC Hussein_NNP Mardan_NNP are_VBP both_

G8 patterns in Non-Native Literature sub-corpus:

	File	Left Context	Hit	Right Context
1	LITERATU...	NNS of_IN the_DT two_CD fathers_NNS reflect_	VBP	the_DT dominance_NN and_CC control_NN over_
2	LITERATU...	DT following_VBG three_CD questions_NNS :_: What_WP are_	VBP	the_DT elements_NNS that_WDT make_VBP Bingo_
3	LITERATU...	PRP$ family_NN in_IN general_NN ?_. What_WP are_	VBP	the_DT real_JJ drives_NNS which_WDT stand_
4	LITERATU...	NNS "_" ,_, meaning_VBG the_DT characters_NNS who_WP represent_	VBP	the_DT struggle_NN of_IN individuals_NNS in_
5	LITERATU...	PRP$ harsh_JJ reality_NN ._. The_DT plays_NNS present_	VBP	the_DT suffering_NN of_IN the_DT two_
6	LITERATU...	VBP authoritative_JJ and_CC dominating_VBG .They_PRP want_	VBP	to_TO control_VB and_CC command_VB the_
7	LITERATU...	NN ._. Nevertheless_RB ,_, the_DT three_CD poets_NNS seek_	VBP	to_TO glorify_VB Jerusalem_NNP 's_NN holiness_
8	LITERATU...	IN modern_JJ poetry_NN ._. This_DT study_NN aim_	VBP	to_TO investigate_VB modern_JJ literature_NN 's_
9	LITERATU...	NNS of_IN women_NNS which_WDT eventually_RB lead_	VBP	to_IN inflicting_VBG some_DT damaging_JJ effects_
10	LITERATU...	TO analyze_VB the_DT elements_NNS that_WDT make_	VBP	Bingo_NNP a_DT biographical_JJ fiction_NN by_
11	LITERATU...	WP are_VBP the_DT elements_NNS that_WDT make_	VBP	Bingo_NNP a_DT biographical_JJ fiction_NN ?_. How_
12	LITERATU...	IN the_DT various_JJ problems_NNS women_NNS face_	VBP	in_IN patriarchal_JJ societies_NNS ._. This_DT study_
13	LITERATU...	DT injustices_VBZ the_DT female_JJ characters_NNS suffer_	VBP	in_IN their_PRP$ man_NN -_HYPH dominated_VBN
14	LITERATU...	VBG these_DT female_JJ characters_NNS that_WDT have_	VBP	many_JJ differences_NNS and_CC gaps_NNS between_

G1 patterns in Native translation sub-corpus:

	File	Left Context	Hit	Right Context
1	Translatio...	NN translation_NN strategy_NN based_VBN on_IN Vinay_	NNP	and_CC Darbelnet_NNP 's_POS (_-LRB- 1995_CD)_-RRB-
2	Translatio...	NN including_VBG formal_JJ lessons_NNS on_IN Vinay_	NNP	and_CC Darbelnet_NNP 's_POS indirect_JJ translation_
3	Translatio...	VBN by_IN King_NNP Vajiravudh_NNP of_IN Siam_	NNP	and_CC it_PRP is_VBZ based_VBN on_
4	Translatio...	VBD translated_VBN from_IN Thai_NNP to_IN English_	NNP ._	and_CC it_PRP was_VBD adapted_VBN into_
5	Translatio...	VB CCG_NNP ._. Spellcheckers_NNS such_JJ as_IN ASPELL_	NNP	and_CC Microsoft_NNP Word_NNP 2007_CD were_VBD
6	Translatio...	NNP RON_NNP ._. ErronLT_NNP R_NN outperforms_VBZ ASPELL_	NNP	and_CC Microsoft_NNP Word_NNP 2007_CD by_IN
7	Translatio...	at_IN a_DT reduced_JJ cost_NN (_-LRB- Brill_	NNP	and_CC Moore_NNP ,_, 2000_CD)_-RRB- ,_, (_-LRB- Toutanova_NNP
8	Translatio...	Brill_NNP and_CC Moore_NNP ,_, 2000_CD)_-RRB- ,_, (_-LRB- Toutanova_	NNP	and_CC Moore_NNP ,_, 2002_CD)_-RRB- ._. Boyd_NNP (_-LRB- 2008_
9	Translatio...	letter_NN -_HYPH based_VBN model_NN ,_, ErronLT_NNP R_	NNP ._	and_CC a_DT pronunciation_NN -_HYPH based_VBN
10	Translatio...	pectives_NNS ._. Schaeffer_NNP ,_, Dragsted_NNP ,_, Hvelplund_NNP ,_, Balling_	NNP ._	and_CC Carl_NNP (_-LRB- 2016b_CD)_-RRB- provided_
11	Translatio...	NN of_IN the_DT Philosophy_NNP of_IN Science_	NNP	and_CC examines_VBZ how_WRB to_TO differentiate_
12	Translatio...	VB up_RP to_IN Communication_NNP Studies_NNPS ,_, Media_	NNP	and_CC Film_NNP Studies_NNPS ,_, Cultural_NNP Studies_
13	Translatio...	I Semiotics_NNP ,_, Sociology_NNP ,_, Anthropology_NNP "_" (_-LRB- Gambier_	NNP	and_CC Gottlieb_NNP xii_NNP)_-RRB- ,_, and_CC
14	Translatio...	IN translations_NNS ._. This_DT study_NN employs_VBZ Morris_	NNP	and_CC Hirst_NNP 's_POS (_-LRB- 1991_CD)_-RRB-

G8 patterns in Native translation sub-corpus:

	File	Left Context	Hit	Right Context
1	Translatio...	NN needs_NNS or_CC expectations_NNS and_CC do_	VBP	not_RB fulfill_VB business_NN objectives_NNS ._. This_
2	Translatio...	IN while_IN JWEFL_JJ spelling_NN errors_NNS do_	VBP	not_RB have_VB the_DT same_JJ patterns_
3	Translatio...	JJ at_IN this_DT task_NN ,_, they_PRP do_	VBP	not_RB perform_VB well_RB when_WRB used_
4	Translatio...	NNP spelling_NN errors_NNS which_WDT simply_RB are_	VBP	not_RB present_JJ in_IN Japanese_JJ learner_
5	Translatio...	syntax_NN -_HYPH based_VBN techniques_NNS that_WDT do_	VBP	not_RB use_VB CCG_NNP ._. Spellcheckers_NNS such_
6	Translatio...	knowledge_NN -_HYPH sharing_VBG tools_NNS that_WDT promote_	VBP	the_DT development_NN of_IN technological_JJ and_
7	Translatio...	SP In_IN this_DT thesis_NN ,_, I_PRP analyze_	VBP	the_DT different_JJ obstacles_NNS of_IN audiovisual_
8	Translatio...	VBG for_IN translation_NN ,_, these_DT findings_NNS suggest_	VBP	the_DT study_NN of_IN words_NNS that_
9	Translatio...	and_CC socio_NN -_HYPH critical_JJ references_NNS are_	VBP	the_DT two_CD main_JJ components_NNS of_
10	Translatio...	NN between_IN how_WRB software_NN designers_NNS envision_	VBP	the_DT use_NN of_IN their_PRP$ application_
11	Translatio...	NNS that_WDT traditional_JJ syntactic_JJ theories_NNS struggle_	VBP	to_TO describe_VB ._. These_DT CCG_NN labels_
12	Translatio...	NNS of_IN tools_NNS professional_JJ translators_NNS employ_	VBP	to_TO interact_VB with_IN other_JJ professionals_
13	Translatio...	VBZ ,_, software_NN tools_NNS and_CC technology_NN fail_	VBP	to_TO meet_VB user_NN needs_NNS or_
14	Translatio...	JJ edition_NN of_IN Scientific_NNP American_NNP ,_, tend_	VBP	to_TO reproduce_VB the_DT lexical_JJ cohesive_

G1 patterns in Non-Native Translation sub-corpus:

	File	Left Context	Hit	Right Context
1	translatio...	CC focus_NN of_IN research_NN about_IN Arabic_	NNP -_	HYPH English_NNP MT_NNP systems_NNS ._. That_DT
2	translatio...	CC translation_NN linguistic_JJ level_NN in_IN Arabic_	NNP -_	HYPH English_NNP MT_NNP studies_NNS ._. The_DT
3	translatio...	NN MT_NNP of_IN studies_NNS about_IN Arabic_	NNP -_	HYPH English_NNP MT_NNP systems_NNS got_VBD
4	translatio...	IN genre_NN on_IN studies_NNS about_IN Arabic_	NNP -_	HYPH English_NNP MT_NNP systems_NNS was_VBD
5	translatio...	NN used_VBN in_IN studies_NNS about_IN Arabic_	NNP -_	HYPH English_NNP MT_NNP systems_NNS was_VBD
6	translatio...	NN direction_NN in_IN studies_NNS about_IN Arabic_	NNP -_	HYPH English_NNP MT_NNP systems_NNS obtained_VBD
7	translatio...	NN in_IN the_DT studies_NNS about_IN Arabic_	NNP -_	HYPH English_NNP MT_NNP systems_NNS ._. As_IN
8	translatio...	frequency_NN in_IN studies_NNS about_IN __SP Arabic_	NNP -_	HYPH English_NNP MT_NNP systems_NNS ._. The_DT
9	translatio...	dictionary_NNP)_-RRB- in_IN studies_NNS about_IN Arabic_	NNP -_	HYPH English_NNP MT_NNP was_VBD BLEU_NNP ._.
10	translatio...	IN feminine_JJ expressions_NNS from_IN Surat_NNP An_	NNP -_	HYPH Nisa_NNP and_CC their_PRP$ translation_NN
11	translatio...	NNPS :_: Translation_NN ,_, Feminine_JJ expressions_NNS ,_, Surat_NNP An_	NNP -_	HYPH Nisa_NNP ._. This_DT study_NN aims_VBZ
12	translatio...	bstacles_NNS ,_, strategies_NNS ,_, translation_NN ,_, proverbs_NNS ,_, English_	NNP -_	HYPH Arabic_NNP This_DT study_NN aims_VBZ
13	translatio...	IN the_DT Arabic_JJ novel_NN Saq_NNP Al_	NNP -_	HYPH Bambu_NNP ,_, and_CC identify_VB the_DT
14	translatio...	IN John_NNP Medows_NNP Rodwell_NNP ,_, Muhammad_NNP Al_	NNP -_	HYPH Hilali_NNP and_CC Khan_NNP ,_, and_CC

G8 patterns in Non-Native Translation sub-corpus:

	File	Left Context	Hit	Right Context
1	translatio...	NNS that_WDT Jordanian_JJ novice_NN translators_NNS encounter_	VBP	when_WRB translating_VBG proverbs_NNS and_CC to_
2	translatio...	TO explore_VB the_DT strategies_NNS they_PRP use_	VBP	when_WRB translating_VBG proverbs_NNS from_IN Arabic_
3	translatio...	NNS which_WDT Jordanian_JJ novice_NN translators_NNS face_	VBP	when_WRB translating_VBG proverbs_NNS from_IN Arabic_
4	translatio...	NNS that_WDT Jordanian_JJ novice_NN translators_NNS face_	VBP	when_WRB they_PRP translate_VBP proverbs_NNS ?_. 2_LS -_
5	translatio...	IN the_DT journalist_NN 's_POS translators_NNS have_	VBP	a_DT high_JJ level_NN in_IN the_
6	translatio...	IN cultural_JJ bias_NN when_WRB they_PRP translate_	VBP	a_DT journalistic_JJ text_NN ._. The_DT researcher_
7	translatio...	JJ results_NNS of_IN character_NN recognition_NN show_	VBP	a_DT rate_NN of_IN 81.82_CD %_NN ,_, the_
8	translatio...	VBD wrong_JJ answers_NNS ._. The_DT results_NNS have_	VBP	also_RB indicated_VBN that_IN nine_CD techniques_
9	translatio...	JJ terms_NNS ._. Furthermore_RB ,_, the_DT results_NNS have_	VBP	also_RB showed_VBN that_IN translators_NNS use_
10	translatio...	DT few_JJ cases_NNS of_IN mistranslation_NN are_	VBP	also_RB spotted_VBN ._. The_DT study_NN recommends_
11	translatio...	DT cultural_JJ bias_NN while_IN they_PRP translate_	VBP	the_DT journalistic_JJ texts_NNS ._. Key_JJ words_
12	translatio...	DT following_VBG questions_NNS :_: 1_LS -_UH What_WP are_	VBP	the_DT obstacles_NNS that_WDT Jordanian_JJ novice_
13	translatio...	DT system_NN ._. The_DT system_NN automatically_RB detect_	VBP	the_DT text_NN in_IN the_DT image_
14	translatio...	also_RB reveals_VBZ that_IN 12_CD procedures_NNS have_	VBP	been_VBN employed_VBN by_IN the_DT translators_