



**A CORPUS-BASED STUDY OF THE ENGLISH
SYNONYMS: HARD, TOUGH, AND DIFFICULT**

BY

AZIN BOOJARZADEH

**AN INDEPENDENT STUDY SUBMITTED IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF ARTS IN ENGLISH LANGUAGE TEACHING
LANGUAGE INSTITUTE
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2024**

**A CORPUS-BASED STUDY OF THE ENGLISH
SYNONYMS: HARD, TOUGH, AND DIFFICULT**

BY

AZIN BOOJARZADEH



**AN INDEPENDENT STUDY SUBMITTED IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF ARTS IN ENGLISH LANGUAGE TEACHING
LANGUAGE INSTITUTE
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2024**

THAMMASAT UNIVERSITY
LANGUAGE INSTITUTE

AN INDEPENDENT STUDY

BY

AZIN BOOJARZADEH

ENTITLED

A CORPUS-BASED STUDY OF THE ENGLISH SYNONYMS: HARD, TOUGH,
AND DIFFICULT

was approved as partial fulfillment of the requirements for
the degree of Master of Arts in English Language Teaching

on July 23, 2025

Chairman


(Associate Professor Supakorn Phoocharoensil, Ph.D.)

Member and Advisor


(Associate Professor Supong Tangkiengsirisin, Ph.D.)

Director


(Associate Professor Supakorn Phoocharoensil, Ph.D.)

Independent Study Title	A CORPUS-BASED STUDY OF THE ENGLISH SYNONYMS: HARD, TOUGH AND DIFFICULT
Author	Azin Boojarzadeh
Degree	Master of Arts
Major Field/Faculty/University	English Language Teaching Language Institute Thammasat University
Independent Study Advisor	Associate Professor Supong Tangkiengsirisin, Ph.D.
Academic Year	2024

ABSTRACT

This corpus-based study investigates the similarities and differences among the English adjectives “hard”, “tough”, and “difficult” by analyzing their noun collocations and semantic preferences. Using data from the Corpus of Contemporary American English (COCA), and online dictionaries, the study identified the top 30 noun collocates for each adjective, which were then analyzed and categorized into semantic fields. The findings revealed that while all three adjectives commonly collocate with nouns “time/s”, “decisions”, “choices”, “questions”, and “situations”, each adjective also exhibits different collocational patterns that reflect its unique usage. “Hard” shows the highest frequency of noun collocates, whereas “tough” and “difficult” have lower frequencies but form strong associations with specific nouns. For semantic preference, “hard” is mainly associated with physical and technical terms, “tough” strongly connects with descriptions of people and struggle, and “difficult” is preferred in cognitive and formal domains. Nonetheless, all three adjectives share certain semantic categories related to Abstract Concepts and Time. These collocational patterns and semantic preferences offer valuable insights that can enhance synonym selection and usage accuracy.

Keywords: Corpus-based analysis, Synonyms, Collocation, Semantic fields, COCA (Corpus of Contemporary American English)

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to everyone who has supported me throughout this journey. To my family and friends, thank you for your patience, encouragement, and belief in me, especially during the most challenging moments of my degree. Your support has meant more than words can say.

I am especially thankful to my advisor, Assoc. Prof. Dr. Supong Tangkiengsirisin, for his thoughtful guidance and steady encouragement throughout this study. I also owe special thanks to Assoc. Prof. Dr. Supakorn Phoocharoensil, whose insightful feedback helped shape the direction of my work. It was through his engaging lectures that I first developed an interest in the topic of linguistics.

To everyone who has contributed in ways big or small—thank you. This study is the result of not just individual effort, but also the kindness, knowledge, and support shared with me along the way.

Azin Boojarzadeh

TABLE OF CONTENTS

	Page
ABSTRACT	(1)
ACKNOWLEDGEMENTS	(3)
LIST OF TABLES	(7)
LIST OF ABBREVIATIONS	(8)
CHAPTER 1 INTRODUCTION	1
1.1 Background of the study	1
1.2 Research objectives	2
1.3 Research questions	3
1.4 Definition and key terms	3
1.4.1 Synonyms	3
1.4.2 Corpus	3
1.4.3 COCA	3
1.4.4 Collocation	3
1.4.5 Dictionary	3
1.4.6 Semantic preference	4
1.5 Scope of the study	4
1.6 Significance of the study	4
1.7 Organization of the study	5
1.7.1 Chapter 1 introduction	5
1.7.2 Chapter 2 review of literature	5
1.7.3 Chapter 3 research methodology	5
1.7.4 Chapter 4 results	5
1.7.5 Chapter 5 discussion, conclusion, and recommendations	5

CHAPTER 2 REVIEW OF LITERATURE	7
2.1 Synonyms	7
2.1.1 Definition of synonym	6
2.1.2 Types of synonyms	8
2.2 Criteria for distinguishing synonyms	9
2.2.1 Collocations	9
2.2.2 Sense of meaning	11
2.3 Corpus linguistics and the study of synonyms	11
2.4 Relevant research	12
CHAPTER 3 RESEARCH METHODOLOGY	16
3.1 Target words	16
3.2 Research instruments	16
3.3 Data collection	17
3.4 Data analysis	17
3.4.1 The analysis of noun collocations	18
3.4.2 The analysis of the semantic preference	18
CHAPTER 4 RESULTS	19
4.1 Collocation	19
4.1.1 “Hard” collocates	19
4.1.2 “Tough” collocates	20
4.1.3 “Difficult” collocates	21
4.2 Semantic preference	22
4.2.1 Semantic preference of “Hard”	23
4.2.2 Semantic preference of “Tough”	25
4.2.3 Semantic preference of “Difficult”	27

CHAPTER 5 DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS	31
5.1 Summary of the finding	31
5.1.1 Noun collocations	31
5.1.2 Semantic preference	32
5.2 Discussion	33
5.2.1 Comparing collocations of adjectives “Hard”, “Tough”, and “Difficult”	33
5.2.2 Comparing semantic preference of adjectives “Hard”, “Tough”, and “Difficult”	35
5.3 Conclusion	37
5.4 Implications	37
5.5 Recommendations	38
REFERENCES	39

LIST OF TABLES

Tables	Page
4.1.1 Top 30 Noun Collocates of “Hard” Ranging by Frequency	19
4.1.2 Top 30 Noun Collocates of “Tough” Ranging by Frequency	20
4.1.3 Top 30 Noun Collocates of “Difficult” Ranging by Frequency	22



LIST OF ABBREVIATIONS

Symbols/Abbreviations	Terms
OALD	Oxford Advanced Learner's Dictionary
CALD	Cambridge Advanced Learner's Dictionary
LDOCE	Longman Dictionary of Contemporary English
LAAD	Longman Advanced American Dictionary
COCA	Corpus of Contemporary American English
ESL	English as a second language

CHAPTER I

INTRODUCTION

This chapter presents the background of the study, research objectives, research questions, definitions of the key terms, scope of the study, limitations of the study, and significance of the study.

1.1 Background of the Study

Without a doubt, English is the present universal language or the lingua franca. This indicates that people who speak various first languages communicate with each other via English. As a matter of fact, English is actually one of the most widely used languages in the world, both as a first and a second language (Harmer, 2007). English is most used as a second language, promoting communication in domains including science, technology, business, and education. Its dominance emphasizes how vital it is to become proficient in every area of the language, such as grammar, pronunciation, and vocabulary in order to communicate successfully in a globalized world.

In the dynamic field of English language learning, vocabulary plays a vital role in developing linguistic competence and effective communication (Khan et al., 2018). For English language learners, mastering vocabulary is important not only to convey ideas but also to understand the subtle nuances of context, usage, and meaning. Danglli & Abazaj (2014) stated that synonyms can enhance language by adding variety and precision. However, synonyms also present one of the biggest obstacles. Despite their seeming interchangeability, synonyms are frequently used differently depending on grammatical, collocational, and semantic considerations. In the study conducted by Zeng (2007), challenges with synonyms arise when words that seem similar, like “*participant*” and “*negotiator*” are used in different contexts, leading to confusion if not applied accurately. With this, a deep understanding of synonyms is important for learners to communicate accurately in various contexts while putting importance on the vocabulary’ grammatical, collocational, and semantic considerations.

Although there were several studies conducted about the similarities and differences of other synonymous words, the researcher has not come across a study that specifically examines the similarities and differences of adjectives: *hard*, *tough*, and *difficult*. This set of synonyms is common in English and can sometimes be used interchangeably when describing something challenging, such as “*It is **difficult** to open the door*” can also be said as “*It is **hard** to open the door*”, and “*I find this test **difficult***” can be changed to “*I find this test **tough***”.

However, in some contexts they differ in the semantic preferences. For instance, in the sentence “*This old bread is difficult*” the appropriate adjective would be *hard*, as it pertains to the physical texture of the bread. Similarly, “*The diamond is known to be **hard***” is correct, but not “*The diamond is known to be **difficult***”, unless you are referring to cutting it. *Tough* also has its own specific uses, like “*He is a **tough** person*” meaning resilient, or “*This meat is **tough***”, meaning hard to chew. However, it would be inappropriate to state “*This old bread is **tough***” or “*The diamond is **tough***”, as this would imply the bread or diamond possesses resilience, which does not align with the intended meaning in these contexts. Each word reflects different semantic preferences and understanding their distinct types of challenges ensures learners select the most accurate word for a given situation. With that, learners who have trouble telling how to use each word in which context correctly are frequently perplexed by these subtle differences.

Hence, this study investigates the similarities and differences of these three synonymous adjectives: “*hard*”, “*tough*”, and “*difficult*”, using a corpus-based approach to identify their contextual distinctions in collocation and semantic preferences. Data for this study is gathered from the Oxford Advanced Learner’s Dictionary (2005), the Longman Dictionary of Contemporary English (2009), and the Corpus of Contemporary American English (COCA), providing a comprehensive foundation for analyzing the usage and contextual nuances of the selected synonyms.

1.2 Research Objectives

1. To investigate the similarities and differences among the synonyms *hard*, *tough*, and *difficult* in terms of common noun collocations.
2. To explore the similarities and differences among the synonyms *hard*, *tough*, and *difficult* in terms of their semantic preferences.

1.3 Research Questions

1. What are the similarities and differences among the synonyms *hard*, *tough*, and *difficult* in terms of their common noun collocations?
2. What are the similarities and differences among the synonyms *hard*, *tough*, and *difficult* in terms of their semantic preferences?

1.4 Definitions of Key Terms

1.4.1 Synonyms

Synonyms refer to words that share similar meaning but are not always the same and are not interchangeable in all possible contexts. In this study, the synonyms being studied are “*hard*”, “*tough*”, and “*difficult*”.

1.4.2 Corpus

A corpus refers to a group of written and spoken texts that have been saved on a computer as electronic texts (O'Keeffe et al., 2007). Additionally, the primary tool for gathering data on the selected adjectives in this study is the Corpus of Contemporary American English (COCA).

1.4.3 COCA

COCA (Corpus of Contemporary American English) is a database that currently has over one billion words. With 485,202 texts from eight different genres, including magazines, TV and movie subtitles, spoken language, fiction, websites, newspapers, blogs, and scholarly journals. In this study, this tool helps researchers find synonyms' similarities and differences through contextual analysis.

1.4.4 Collocations

Collocation refers to the nouns that combine with the three target adjectives: “*hard*”, “*tough*”, and “*difficult*”.

1.4.5 Dictionary

A dictionary is a reference source that provides information about words, including their meanings, pronunciation, usage, and grammatical features. In this study, the Longman Dictionary of Contemporary English and Merriam-Webster Dictionary are used as key resources to analyze the definitions, collocations, and contextual applications of “*hard*”, “*tough*”, and “*difficult*”.

1.4.6 Semantic Preference

This refers to the tendency of a word to co-occur with nouns from particular semantic fields or categories. It highlights the kinds of meanings or concepts a word is typically associated with in context. In this study, semantic preference is used to classify the top noun collocates of *hard*, *tough*, and *difficult* into categories such as physical effort, emotional experience, decision-making, abstract ideas, and technical challenges. This classification helps uncover how each adjective is characteristically applied in real-world usage.

1.5 Scope of the Study

This study focuses on the three synonymous adjectives: “*hard*”, “*tough*”, and “*difficult*”. The data for analysis is retrieved from the Longman Dictionary of Contemporary English (2009), Oxford Advanced Learner’s Dictionary (2005), and the Corpus of Contemporary American English (COCA). The study examines the similarities and differences of these three synonyms in terms of common noun collocations and their semantic preferences.

1.6 Significance of the Study

Since vocabulary is an important part for the use of a language, it is important for language users to choose the appropriate vocabulary and put them into the proper context. Despite their frequent interchangeability, the adjectives “*hard*”, “*tough*”, and “*difficult*” have subtle differences in meaning and usage. English language learners will benefit from this study since it will help them comprehend the proper usage of these synonyms in a variety of circumstances, which will enhance their vocabulary and communication abilities in general.

1.7 Organization of the Study

The study consists of five chapters as follows:

1.7.1 Chapter 1 Introduction

This chapter introduces the study, outlining the background, research questions, objectives, definitions of key terms, the scope of the study, the significance of the research, and the structure of the paper. It lays the groundwork for examining the meaning, usage, and collocational similarities and comparison between the synonyms “*hard*”, “*tough*”, and “*difficult*”.

1.7.2 Chapter 2 Review of literature

This chapter offers a comprehensive review of relevant literature, discussing key concepts such as corpus linguistics, theories of synonymy, and previous studies related to the contextual analysis of synonyms. A theoretical foundation for comprehending the differences between these adjectives is developed in this chapter.

1.7.3 Chapter 3 Research Methodology

This chapter outlines the research methodology, detailing the approach used to collect and analyze data from sources like the Oxford Advanced Learner's Dictionary, Longman Dictionary of Contemporary English, and the Corpus of Contemporary American English (COCA). The methods for examining the meaning, usages, and collocations of “*hard*”, “*tough*”, and “*difficult*” are described in this section.

1.7.4 Chapter 4 Results

This chapter presents the results of the data analysis, showcasing the findings based on the corpus data. It offers insights into the semantic, grammatical, and collocational differences among the three adjectives. The data was analyzed by applying the criteria and theories to distinguish each synonym.

1.7.5 Chapter 5 Discussion, Conclusion, and Recommendations

This chapter concludes the study with a summary of the findings, a discussion of the results, and a consideration of the study's limitations. It also underlines the

pedagogical implications for language learners and offers suggestions for additional research, highlighting the ways in which these findings might be used in English language instruction.



CHAPTER 2

REVIEW OF LITERATURE

This chapter reviews some literature and studies related to the present study.

2.1 Synonyms

2.1.1 Definition of Synonym

Synonyms are words that have the same or nearly the same meaning as other words in the same language. For example, “*tall*” and “*high*” are often used to show this idea. According to the Oxford Advanced Learner’s Dictionary (2015), a synonym is a word or phrase with a similar meaning to another, while the Longman Dictionary (2014) also defines synonyms as words with the same meaning.

These kinds of words are found everywhere, it might be found in daily conversations, textbooks, and academic writing. The wide use of synonyms helps language users, especially English learners, to choose the right words for different situations and express themselves more clearly. This variety also adds depth and flexibility to the language (Sakulratanacharoen, 2012).

The idea of synonymy, or the relationship between words with similar meanings, comes from the Greek word *sunonumon*, which means “*having the same name*” (Online Etymology Dictionary, 2016). It is a key topic in lexicology, the study of words and their meanings.

However, research has shown that most synonyms are not exactly the same. Synonyms often have small differences in meaning, grammar, or how they are used in real situations. Because of these differences, many linguists believe that true or perfect synonyms are very rare, or may not exist at all (Taylor, 2002; Chalker & Weiner, 1998). With all this in mind, it can be said that synonyms play a vital role in linguistics, enriching the language and enhancing its versatility. While synonyms convey similar meanings, the subtle differences in their usage, meaning, and context emphasize the importance of distinguishing between various types of synonyms. As a result, linguists have developed a framework to categorize synonyms into different types. Each serves distinct functions in language use.

2.1.2 Types of Synonyms

Synonyms can be grouped into different types. However, this study focuses on two main types of synonyms, these are: strict (or absolute) synonyms and loose (or relative) synonyms. Strict synonyms are words that have exactly the same meaning and can be used in place of each other in any situation. When one word replaces the other, there should be no change in meaning, tone, or feeling in the sentence (Jackson & Amvela, 2000). In such cases, the choice of which word to use depends only on the speaker or writer, because both words fit equally well in the same context. These kinds of synonyms are rare, and when they do occur, they allow full interchangeability without affecting the message. Nevertheless, some linguists argue that slight changes of word meaning can occur. For example, the words “*clear*” and “*transparent*” both refer to something that allows light to pass through, but their connotations differ. “*Clear*” is often used in a positive context, such as clear “*water*” or “*clear skies*”, implying purity and brightness. On the other hand, “*transparent*” can sometimes suggest a lack of depth or complexity, as in “*transparent motives*” or “*transparent material*”. While the two words are similar in meaning, the nuances in their usage affect how they are perceived in different contexts. In this light, McCarthy, O’Keefe and Walsh (2010) brought to realization that truly strict synonyms are remarkably rare or even inexistent.

Loose or relative synonyms are words that share similar meanings but cannot always be used in the same situations (Jackson & Amvela, 2000). Although these words have overlapping meanings, they are not fully interchangeable in all contexts. This type of synonymy is much more common in English than strict synonymy. In many cases, relative synonyms differ slightly in tone, usage, or the type of situation in which they are appropriate. These differences may be related to formality, emotion, or specific subject areas. Therefore, choosing the correct word often depends on the context and the speaker or writer’s intention. For example, the words “*find*” and “*discover*” can replace each other in a sentence like “*Lydia found/discovered the ball behind the garden shed.*” However, they are not always exchangeable, for instance, in the sentence “*Marie Curie discovered radium in 1898,*” the word find would not be suitable (Jackson & Amvela, 2000).

The example of loose synonyms above indicates that synonyms may be used interchangeably when their meanings overlap; they are not always interchangeable,

particularly in more formal or specialized contexts like scientific discoveries. For instance, “*discover*” is used in the historical context of Marie Curie's work with radium because it implies a significant, formal revelation, whereas “*find*” would not convey the same depth of discovery in this context. Unsurprisingly, it is not an easy task for L2 learners to make distinctions of the two synonymous words as they share an eminently close sense of meaning.

2.2 Criteria for Distinguishing Synonyms

A number of characteristics, including dialect, degree of formality, connotation, chunk, collocate, and so on, can be used to differentiate between synonyms. These factors include, but are not limited to, chunk, which refers to the fixed expressions or word combinations in which a synonym might occur more naturally; dialect, where certain synonyms may be more common in specific regional or social varieties of a language; formality degree, where words may differ in their level of formality and are therefore suitable for different social or professional settings; connotation, where synonyms may carry different emotional or cultural associations, influencing their tone and implied meaning; and collocation, which looks at the tendency of words to combine with specific nouns, verbs, or adjectives in ways that affect their meaning and usage. Each of these aspects helps to shape the way in which synonyms are understood and employed in language, ensuring that speakers can select the most appropriate word for their intended meaning. However, in this study, the synonyms were distinguished by closely examining their collocations and semantic preferences.

2.2.1 Collocations

According to the Oxford Advanced Learner's Dictionary (2015), collocations refer to word combinations that frequently occur together in a language, far more often than would be expected by random chance. Lewis (2000) further explains that collocations are not merely words placed side by side but rather patterns where specific words tend to appear together consistently and significantly within real language use. These patterns are driven by natural language tendencies rather than arbitrary selection. For instance, the adjective *auburn* commonly appears with noun *hair* but not with unrelated nouns like *car*, and the adjective *rancid* is typically found alongside *butter*,

rather than with bread (McCarthy, O’Keeffe, & Walsh, 2010). These examples illustrate how collocations reflect statistically preferred pairings that native speakers intuitively recognize and use.

According to Hill (2000), collocations can be divided into four categories: unique collocations, strong collocations, weak collocations, and medium-strength collocations. First, unique collocations, these collocations are fixed and cannot be substituted with alternative words or phrases without changing the meaning or rendering the expression incorrect. These collocations often involve idiomatic phrases or fixed expressions that are ingrained in the language. For example, “*foot the bill*” refers to paying for something, but this phrase cannot be substituted with “*foot the invoice*” without altering the intended meaning. Unique collocations are often cultural and context-dependent, making them highly specific in their usage. They are usually learned through exposure to the language and are not easily deduced from the meanings of the individual words involved.

Second, strong collocations, these collocations are combinations of words that occur together so frequently that the meaning of one word is nearly inseparable from the other in most contexts. These combinations are highly specific and often represent concepts that have a fixed way of expressing themselves. For example, “*rancid oil*” is a strong collocation because the adjective *rancid* almost exclusively co-occurs with *oil*, not with other substances, as the word “*rancid*” carries a strong, specific meaning when paired with oils or fats. Strong collocations have a higher degree of predictability in their use, and their components are almost never separated without causing awkwardness or confusion.

Third, weak collocations, these collocations are pairs of words that can easily be combined with a wide variety of other words without significant restrictions. These collocations are less fixed and have a broad range of flexibility in their usage. Words involved in weak collocations tend to have more literal or generic meanings. For example, “*red shirt*” is a weak collocation because both *red* and *shirt* are commonly used in many other contexts. The word *red* can easily pair with other nouns like “*red apple*”, “*red car*”, or “*red light*”, and *shirt* can be combined with different adjectives like “*cotton shirt*”, “*blue shirt*”, or “*plain shirt*”. Weak collocations are more flexible and typically do not carry as strong a sense of restriction as strong collocations.

Lastly, medium-strength collocations, these collocations represent a balanced category between strong and weak collocations. They are not as rigid as strong collocations, but they are more predictable and commonly used than weak collocations. These combinations are useful in both spoken and written communication, and they form an essential part of the language learning process. For example, the collocation “*hold a meeting*” is not as strictly tied as “*rancid oil*”, but it is a well-established phrase that learners should know. The words “*hold*” and “*meeting*” are frequently paired, and their combined meaning is understood in the context of organizing or conducting a meeting. Medium-strength collocations are often taught to learners as they are key to fluency in natural, everyday communication. They offer a good balance of predictability and flexibility, helping speakers and writers express ideas clearly and naturally.

2.2.2 Sense of Meaning

When understanding a word's meaning, O'Keeffe, McCarthy, and Carter (2007) stress the need of taking into account both the word's connotation and the context in which it appears. When examining the intensity, subtlety, or differing semantic preferences that words may convey in written or spoken conversation, this method is very helpful. For instance, a challenge might involve physical effort or resistance, such as completing a hard manual task; mental complexity, such as solving a difficult problem; or emotional difficulty, such as making a tough decision. Each type of challenge brings its own nuances to the use of language, which is why understanding the broader context is essential for accurate interpretation.

Furthermore, according to Cruse (2010), a word's similarity typically indicates its expressive meaning and can be a trustworthy way to identify difference. Additionally, Murphy (2003) points out that it is common for words to have many meanings and for their denotative meaning to be connected to one of those meanings.

2.3 Corpus Linguistics and the Study of Synonyms

Corpus linguistics is a research methodology that involves examining a language through its authentic use in various contexts (Lindquist, 2009). This approach is particularly valuable or vital in the study of synonyms, as it enables researchers to

analyze how words are used in real-world situations. The methodology utilizes computers to conduct quantitative analyses, providing insights into language features such as grammar, vocabulary, and usage patterns. Moreover, corpus linguistics plays a significant role in fields such as discourse analysis and second language acquisition (O’Keeffe, McCarthy, & McCarthy, 2007).

Moreover, according to Moon (2010), corpora are powerful tools for distinguishing between synonymous words by examining various linguistic dimensions, including genre, frequency, and collocational patterns. Similarly, Bennett (2010) highlights that corpus linguistics helps uncover the associations and patterns linked to lexical and grammatical features of language. By leveraging these tools, researchers can gain a deeper understanding of how synonyms function in different contexts.

Through the use of corpus linguistics, this study aims to explore how these synonymous adjectives differ not only in meaning but also in terms of the challenges they describe and the specific noun phrases they typically co-occur with. This approach will provide a more comprehensive understanding of the subtle distinctions that exist between these words.

Furthermore, a substantial and expanding corpus of research has investigated synonym differences using corpus-based data gathered from diverse corpora, such as the Corpus of Contemporary American English (COCA). This indicates that the corpus-based method to language analysis is a trustworthy source of empirical data. Most academics utilize standard criteria, such meanings, collocations, grammatical patterns, and degree of formality, to distinguish between synonyms. This method would assist language instructors and students in identifying variations in language usage. Furthermore, while some research employed dictionaries as an additional data source, the bulk of the studies solely looked at corpus-based data.

2.4 Relevant Research

Numerous previous studies investigated synonym analysis using corpus evidence from various corpora, such as the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC). Collocations, grammatical patterns, and meanings and senses of meanings were the most often utilized criteria by the majority of studies to distinguish between synonyms.

Mengyun Chen (2023) conducted a corpus-based study examining the similarities and differences of three adjective synonyms—“*funny*”, “*humorous*”, and “*hilarious*”—using two primary criteria: degree of formality and collocations. The study utilized the Longman Dictionary of Contemporary English for meaning extraction and the Corpus of Contemporary American English (COCA) for gathering data from eight different genres. The results revealed that all three adjectives are primarily used in informal contexts, indicating a low degree of formality. Regarding their collocations, while the three adjectives shared several common nouns and adverb collocates, each word also demonstrated unique collocates. This suggests that the three synonyms cannot be used interchangeably in all contexts, as their collocational preferences and nuances vary.

Similarly, Aroonmanakun (2015) used a corpus-based technique to investigate the English adjective synonyms “*quick*” and “*fast*”. The study concentrated on noun collocations to highlight the distinctions and similarities between the two terms by examining COCA and contrasting the fundamental meanings of these synonyms using the Longman Dictionary of Contemporary English. Although the meanings of the two terms overlap, the corpus data demonstrated significant differences in their collocational patterns, according to the findings. It was discovered that “*fast*” concentrated on the “*manner of action or motion*”, but “*quick*” emphasized the “*quality of action*”. For example, “*quick learner*” refers to someone who learns easily, whereas “*fast learner*” denotes someone who learns within a short period. This study highlights the importance of corpus data in understanding the nuances of synonym use, as dictionary definitions alone may not fully capture the contextual differences between these words.

Pantong Imprapai’s (2023) research investigated the near-synonyms *trip*, *journey*, and *voyage*, with particular attention to their nuanced contextual distinctions. The study incorporated definitions from authoritative sources such as the Oxford Advanced Learner’s Dictionary and the Longman Dictionary of Contemporary English, alongside an analysis of usage data drawn from the Corpus of Contemporary American English (COCA). The findings revealed that while these terms are often regarded as interchangeable, they exhibit distinct preferences for use in specific genres and contexts. “*Trip*” is primarily used in informal, everyday situations, while “*journey*” and “*voyage*” are more common in formal and literary contexts. The study also highlighted the specific

verb and adjective collocates associated with each term, emphasizing their unique semantic and pragmatic features. This research provides valuable insights for English language learners and educators, offering a deeper understanding of how to choose the most appropriate synonym for effective communication.

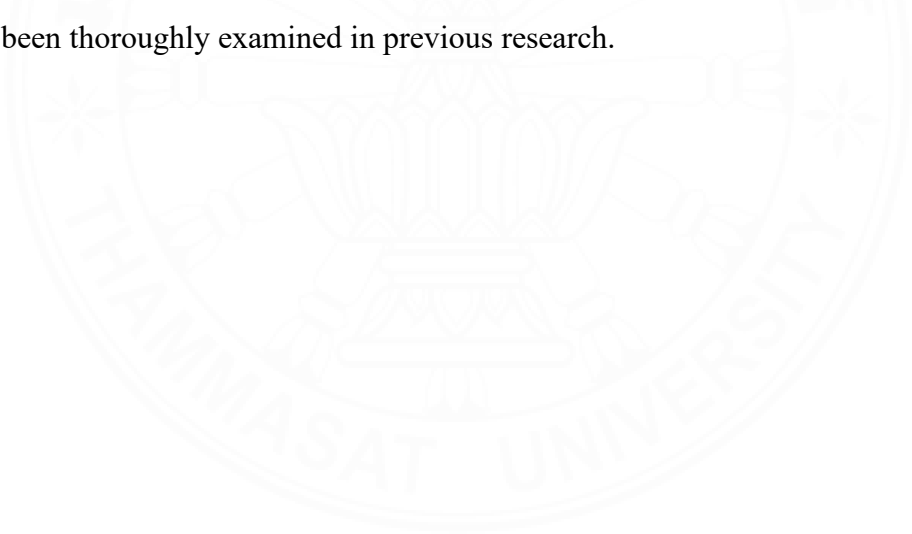
Phoocharoensil (2010) studied the five English synonyms "*ask*", "*beg*", "*plead*", "*request*", and "*appeal*" using corpus-based research, focusing on lexical, syntactic, and stylistic details. By comparing fundamental data from three dictionaries. The data for study was then obtained from a corpus of Time (1995) using Wordsmith Tools 3.0. The study revealed that while the five verbs share a basic meaning related to making a request, they differ in nuance. "*Beg*" conveys urgency, "*plead*" implies emotional intensity, "*request*" reflects politeness and formality, and "*appeal*" suggests seriousness and is often found in legal contexts. Analysis of concordance lines for request also uncovered usage details not noted in the dictionaries consulted.

The degree of formality and noun and adverb collocations with semantic preference were the main topics of Oraya Sriyang's (2023) study, which compared and contrasted the adjectives "rich," "wealthy," and "affluent." The study investigated these synonyms using the Corpus of Contemporary American English (COCA) and online dictionaries Merriam-Webster and the Longman Dictionary of Contemporary English. The findings revealed that "*affluent*" is the most formal, followed by "*wealthy*", with "*rich*" being the least formal. In terms of collocations, while all three adjectives are used to describe people, occupations, and money, they show distinct semantic preferences. For example, "*rich*" is often used in the context of food, facts, and events, while "*affluent*" and "*wealthy*" are used to describe people and locations. The study concluded that these adjectives are near-synonyms but cannot be used interchangeably in all contexts due to differences in formality and collocational behavior.

Cai (2012) studied great and its six synonyms: "*awesome*", "*excellent*", "*fabulous*", "*fantastic*", "*great*", and "*wonderful*". The data was drawn from the Corpus of Contemporary American English (COCA). The findings show that the word "great" is used more often than the other synonyms. Compared to the others, the meaning of "*great*" is weaker and more comprehensive. The words "*fabulous*", "*fantastic*", "*great*" and "*wonderful*" appear to be more frequently used in spoken language, but "*awesome*" and "*excellent*" are more frequently used in periodicals. These synonyms have some

overlapping noun collocates and varying numbers of adverb collocates.

Thus, these previous studies on synonym analysis using a corpus based approaches have revealed significant insights into the fine distinctions of synonymous words. For example, Mengyun Chen (2023) examined adjectives like “*funny*”, “*humorous*”, and “*hilarious*”, highlighting their distinct collocational preferences and contexts. Similarly, Aroonmanakun (2015) analyzed “*quick*” and “*fast*”, finding variations in collocations that pointed to their nuanced differences in meaning. Study by Pantong Imprapai (2023) and Phoocharoensil (2010) further demonstrated how even synonyms with overlapping meanings, like “*trip*”, “*journey*”, “*voyage*”, and “*ask*”, “*beg*”, “*plead*”, differ based on formality, urgency, and context. These studies highlight how vital it is to investigate synonyms in context in order to completely comprehend how they are used. Building on this body of work, the present study investigates the contextual similarities and differences among “*hard*”, “*tough*”, and “*difficult*”, focusing on the types of challenges they represent and their collocational patterns, which have not been thoroughly examined in previous research.



CHAPTER 3

RESEARCH METHODOLOGY

This chapter presents the methodology used in this study. It includes target words, research instruments, and data analysis.

3.1 Target Words

Three terms with comparable meanings, namely: “*hard*”, “*tough*”, and “*difficult*”, were the focus of this study. These terms were chosen because they are often used in written and spoken English, and they are applicable to a variety of difficulties. Also, Longman Communication 3000 introduced the three adjectives. Hence, this suggested that in order to communicate effectively both in writing and speaking, it was crucial to learn these three synonyms.

Despite their similarity, there has been little detailed academic exploration into the differences between these adjectives, especially regarding their contextual usage. Therefore, this study aims to analyze the similarities and differences among “*hard*”, “*tough*”, and “*difficult*” by examining their collocational patterns across various genres and their semantic preferences, thus providing a more evidence-based understanding of these commonly used adjectives is vital.

3.2 Research Instruments

Online dictionaries and corpora are the two primary research tools utilized in this study.

Longman Dictionary of Contemporary English (LDOCE) and Merriam Webster Dictionary were selected because they are well-known and reliable tools. The dictionaries do not only provide meanings, pronunciation, and example sentences like general dictionaries, but they also present synonyms, antonyms, collocations, and related words.

The Corpus of Contemporary American English (COCA) consisted of more than one billion words of text. The corpus presented eight genres: spoken, fiction, popular magazines, newspapers, academic texts, TV and Movies subtitles, blogs, and

other web pages.

The research chose to examine COCA for two primary reasons: the vast volume of material it contains and the regularity with which new texts are added to the corpus. First, compared to the British National Corpus (BNC), the Corpus of Contemporary American English is five to six times bigger. The corpus is also dynamic. It indicates that in order to expand the corpus, current English texts have been uploaded to COCA.

3.3 Data Collection

The data collection process involved retrieving lexical and collocational information for the adjectives “*hard*”, “*tough*”, and “*difficult*” from both corpus and dictionary sources. Online dictionaries such as Longman Dictionary of Contemporary English (LDOCE) and Merriam-Webster were consulted to obtain precise definitions, example sentences, collocations, and semantic context of each target word. These dictionaries provided an initial understanding of meaning and usage.

Using the Corpus of Contemporary American English (COCA) the corpus data was collected, accessed via the official COCA website. The advanced search functions of the corpus interface were used to retrieve noun collocates for each target adjective. The collocation analysis used was to set and sorted the collocates by frequency and mutual information (MI) scores to ensure statistical relevance and contextual salience.

For each adjective, the top 30 noun collocates were extracted and documented. These collocates were then verified manually in the corpus by reviewing sample concordance lines to confirm consistent and meaningful co-occurrence. The data were exported and organized into tables for comparative analysis. This process allowed the study to focus not only on frequency data but also on semantic patterns and contextual usage across different genres, which are vital for understanding the subtle distinctions among the three synonyms.

3.4 Data Analysis

This study aims to address two research questions by analyzing the common noun collocations and the type of challenge represented by the synonymous adjectives “*hard*”, “*tough*”, and “*difficult*”. The analysis focused on identifying the top thirty (30) noun collocates for each adjective using corpus data. Also, the study categorized these

collocations and their semantic preferences. This approach provides a clearer understanding of how these adjectives differ in usage and meaning.

3.4.1 The Analysis of Noun Collocations

The Corpus of Contemporary American English (COCA) provided the data used in this investigation. The study identified and analyzed the top thirty (30) most frequent noun collocates for the “*hard*”, “*tough*”, and “*difficult*” adjectives. To determine the most common and significant noun collocations, frequency analysis was conducted, selecting thirty (30) nouns that most frequently co-occur with each adjective. Additionally, corpus data was examined to provide contextual examples illustrating how these adjectives are used with their collocates. This approach helped highlight both shared and distinctive collocational patterns among the three synonyms.

3.4.2 The Analysis of the Semantic Preference

The identified and analyzed noun collocations were classified and analyzed from their concordance lines from COCA into categories based on their semantic preferences. Also, to support accurate categorization, definitions and usage examples from Merriam-Webster Dictionary and the Longman Dictionary of Contemporary English (LDOCE) were also used to confirm the definitions and subtleties of each word collocate. This classification allows us to observe the kinds of concepts or contexts that “*hard*”, “*tough*”, and “*difficult*” are most frequently associated with. For instance, “*hard*” often collocates with nouns related to effort or tangible outcomes, such as “*work*”, “*labor*”, and “*exercise*”. “*Tough*” tends to appear with terms like “*decision*”, “*experience*”, and “*situation*”, reflecting its use in more situational or evaluative contexts. Meanwhile, “*difficult*” frequently combines with more abstract or conceptual nouns like “*problem*”, “*theory*”, and “*concept*”.

CHAPTER 4

RESULTS

4.1 Collocation

Since all the words studied in this research are adjectives, the primary collocates that this study will examine are the nouns they modify. As previously mentioned, “*hard*”, “*tough*”, and “*difficult*” are close synonyms. Therefore, it is likely that they share some common noun collocates while also exhibiting distinct preferences depending on context. In this study, researcher use the Corpus of Contemporary American English (COCA) to identify the top 30 most frequent noun collocates for each adjective, considering only those with a Mutual Information (MI) score of 3.0 or higher. Furthermore, because researcher is interested in the specific types of nouns these adjectives tend to describe, researcher focused exclusively on nouns that appear after the studied adjectives.

4.1.1 “*Hard*” Collocates

To better understand the typical noun collocates of the adjective “*hard*”, this study extracted the top 30 most frequent nouns that co-occur with hard using data from the Corpus of Contemporary American English (COCA).

Table 4.1.1

Top 30 Noun Collocates of “Hard” Ranging by Frequency

Collocates	Frequency	MI Value	Collocates	Frequency	MI Value
work	11813	4.55	Time	11108	3.40
drive	3736	5.64	Times	3129	3.83
feelings	1125	5.14	Evidence	1108	3.47
drives	1036	6.20	Rock	994	4.19
disk	954	7.13	Core	726	4.55
currency	684	5.97	Labor	664	3.95
worker	661	5.23	Choices	655	4.64

copy	609	4.61	Facts	488	3.86
decisions	424	3.37	Hearing	371	3.05
knocks	352	7.16	Hat	346	4.03
cash	334	3.16	Liquor	256	5.63
candy	250	4.28	Plastic	241	3.15
hats	229	5.27	Lessons	215	3.47
cafe	199	4.83	Truths	193	5.48
lesson	182	3.06	Sciences	181	3.85

Table 4.3.1 shows these nouns collocates arranged by frequency, along with their respective MI scores. Results show that the adjective “*hard*” frequently co-occurs with nouns associated with effort, physical material, emotional intensity, and cognitive challenges. In fact, high frequency collocates such as “*work*” (11,813) and “*time*” (11,108) demonstrate the common and expected usage of *hard* in everyday contexts. In contrast, collocates such as “*disk*” (MI 7.13), “*knocks*” (MI 7.16), and “*drives*” (MI 6.20) exhibit particularly high MI scores despite lower raw frequencies, suggesting a strong but more specific lexical relationship with “*hard*”. These values highlight how some noun collocates are highly typical in specific expressions or technical contexts, even if they occur less often overall.

4.1.2 “*Tough*” Collocates

In this section, the 30 most frequent noun collocates of the adjective “*tough*” were retrieved from the Corpus of Contemporary American English (COCA) to provide a clearer understanding of the typical contexts in which this adjective is used.

Table 4.1.2

Top 30 Noun Collocates of “Tough” Ranging by Frequency

Collocates	Frequency	MI Value	Collocates	Frequency	MI Value
time	2087	3.03	Times	1915	5.16
guy	1780	5.47	Questions	1780	5.47
decisions	868	6.44	Love	783	3.38

job	679	3.84	Decision	605	4.94
choices	591	6.53	Question	591	3.47
guys	458	3.74	Call	453	3.01
situation	433	4.38	Spot	365	5.23
issues	322	3.60	Luck	304	5.30
road	287	3.67	Sell	280	4.56
crime	270	4.42	Fight	246	3.65
competition	231	4.81	Stuff	219	3.24
break	218	3.30	Choice	206	3.53
nails	201	7.00	Economy	200	3.38
loss	194	3.81	Kid	190	3.51
position	183	3.09	Battle	165	4.03

The table above presents the top 30 noun collocates of the adjective “tough”, ranked by frequency, along with their corresponding MI scores. The results indicate that “tough” frequently co-occurs with nouns related to decision-making, personal challenges, and emotional or situational difficulties. High-frequency collocates such as “time” (2,087) and “guy” (1,780) exemplify the common use of tough in everyday expressions. In contrast, nouns like “nails” (MI 7.00), “questions” (MI 5.47), and “choices” (MI 6.53) exhibit high MI values, despite lower raw frequencies. These elevated MI scores suggest that these noun collocates have a stronger, more specialized connection with “tough”, typically appearing in specific contexts or phrases. This pattern illustrates how certain nouns are more strongly associated with “tough” in particular domains.

4.1.3 “Difficult” Collocates

This table presents the top 30 noun collocates of the adjective “difficult”, arranged by frequency, along with their corresponding MI scores.

Table 4.1.3*Top 30 Noun Collocates of “Difficult” Ranging by Frequency*

Collocates	Frequency	MI Value	Collocates	Frequency	MI Value
times	1301	3.74	Task	1160	6.14
situation	1151	4.93	Decision	733	4.35
Decisions	641	5.14	Position	609	3.96
Questions	607	3.63	Circumstances	466	5.33
Issues	462	3.26	Situations	419	5.48
choices	338	4.86	Period	328	3.30
challenge	294	3.78	Conditions	292	3.62
challenges	215	4.16	Tasks	203	4.99
transition	192	4.45	Journey	155	3.95
Gauge	139	6.20	Terrain	137	5.65
moments	125	3.30	Concept	124	3.06
conversations	104	4.03	Proposition	85	4.67
concepts	77	3.63	Aspect	73	73
negotiations	66	3.43	Balancing	60	4.88
topics	60	3.29	Adjustment	53	3.93

The data reveals that the adjective “*difficult*” is frequently used with nouns pertaining to decision-making, situations, and abstract or cognitive challenges. Common high frequency collocates such as “*times*” (1,301), “*situation*” (1,151), and “*task*” (1,160) emphasize the broad and general use of the adjective. On the other hand, nouns like “*gauge*” (MI 6.20), “*terrain*” (MI 5.65), and “*situations*” (MI 5.48) showed a high MI scores, indicating a stronger lexical association with “*difficult*”, often within more specific or technical contexts. This suggests that while “*difficult*” is widely applicable, it also forms particularly strong collocations in specialized expressions.

4.2 Semantic Preference

This part of the study investigates the semantic preferences of the adjectives “*hard*”, “*tough*”, and “*difficult*” by categorizing their top noun collocates according to shared meanings or conceptual fields. Semantic preference refers to the tendency of a word to co-occur with items from particular semantic categories, which can reveal deeper insights into how each adjective is typically used in context. To do this, the top 30 noun collocates identified in the previous section, each with an MI score of 3.0 or higher—were grouped into categories based on dictionary definitions, contextual usage, and concordance line analysis. By examining these categories, this paper aims to highlight the subtle distinctions in meaning and usage between the three adjectives, despite their overall semantic similarity.

4.2.1 Semantic Preference of “*Hard*”

In an effort to explore the collocates of “*hard*” in greater depth, the top 30 noun collocates from Table 4.3.1 were grouped into semantic categories based on their shared meanings and usage contexts. Dictionary definitions and concordance lines from COCA were consulted to ensure accurate classification. The collocates were then organized into Eight (8) main categories as follows:

Category	Nouns
1. Abstract Concepts	work (11,813), feelings (1,125), choices (655), decisions (424), lessons (215), lesson (182), hearing (371), truths (193), facts (488)
2. Time	time/times (14,237)
3. Object	drive/drives (4,772), disk (954), copy (609), hat (346), hats (229), candy (250), plastic (241), liquor (256)
4. Economy	currency (684), labor (664), cash (334)
5. People	worker (661)
6. Events	knocks (352)
7. Places	cafe (199)
8. Knowledge	sciences (181)

Results revealed that the adjective “*hard*” exhibits a strong semantic

preference for collocating with nouns that represent abstract concepts, time, and objects, as well as those linked to economic terms, people, and events. Based on the top 30 noun collocates retrieved from COCA and classified into eight categories, guided by dictionary meanings and contextual evidence from concordance lines on the usage of *“hard”* reflects both literal and metaphorical applications.

The most dominant category is Abstract Concepts, which includes high-frequency collocates such as *“work”*, *“feelings”*, *“choices”*, *“decisions”*, *“lessons”*, *“lesson”*, *“hearing”*, *“truths”*, and *“facts”*. These nouns emphasize mental, emotional, or moral exertion. For instance, *“hard work”*, *“hard choices”*, and *“hard decisions”* all convey sustained effort or complexity in decision-making and emotional experience.

The Time category, represented by *“time/times”*, also shows high frequency. In phrases like *“hard times”*, the adjective intensifies the adversity associated with a temporal period, emphasizing hardship endured over a duration.

Collocates in the Objects category, such as *“drive/drives”*, *“disk”*, *“copy”*, *“hat/hats”*, *“candy”*, *“plastic”*, and *“liquor”* point to the literal meaning of hard as physically solid or resistant. These often appear in technical or material contexts, e.g., *“hard disk”* or *“hard plastic”*, where the adjective emphasizes tangibility and firmness. Within the Economy category, collocates like *“currency”*, *“labor”*, and *“cash”* suggest financial or socio-economic intensity. For example, *“hard currency”* refers to stable, widely accepted money, while *“hard labor”* connotes physically demanding or penal work. These collocates reinforce hard as denoting intensity, difficulty, or inflexibility in economic contexts.

The People category includes *“worker”*, where hard often signifies diligence or resilience, as in *“hard worker”*. This aligns with the broader theme of effort or perseverance in labor.

Events, represented by *“knocks”*, reflect metaphorical adversity. The idiom *“hard knocks”* suggests life’s harsh lessons or difficult experiences that build character through resilience.

The remaining two categories, Places *“cafe”* and Knowledge *“sciences”* indicate more specific contexts. For example, *“hard sciences”* distinguishes empirically rigorous fields like physics and chemistry from the humanities. There are

some examples of this group of semantics relating to *hard* found by retrieving the data from concordance lines of the COCA, which are presented as follows in (1), and (2).

(1) *Thank you so much for all the **hard** work, sacrifice, and the love you showed to all your followers bb.*

(2) *Samsung has learned **hard** lessons in the past about launching devices that weren't ready for prime time.*

In excerpt (1), “*hard work*” is placed in the Abstract Concepts category because it refers to the general idea or quality of effort, rather than a specific, countable task or physical object. Although “work” may involve physical or tangible activities, the phrase “hard work” emphasizes the enduring effort, determination, and commitment behind those actions, qualities that are intangible and conceptual.

In excerpt (2), “*hard lessons*” refers to challenging experiences from which Samsung has learned. This usage indicates the lessons which resulted from adverse situations require reflection and resilience. It also belongs to the Abstract Concepts category, as “lessons” denotes conceptual knowledge gained through difficulty, reinforcing hard as a marker of struggle and growth.

Overall, the findings indicate that “hard” frequently collocates with nouns expressing effort, challenge, endurance, and resistance, whether in emotional, cognitive, material, temporal, or economic domains. The repeated use of hard in such collocations reflects its semantic versatility across both concrete and abstract registers.

4.2.2 Semantic Preference of “Tough”

For the semantic preference of the adjective “*tough*”, the top 30 noun collocates from Table 4.1.2 were grouped into semantic categories based on their meaning and usage patterns. Dictionary entries and concordance line analyses were consulted to ensure accurate classification. The collocates were categorized into Seven (7) main groups, as outlined below:

Category	Nouns
1. Abstract Concepts	question/questions (2,371), decisions/decision (1,473), choices/choice (797), love (783), job (679), situation (433), issues (322), luck (304), loss (194), position (183)

2. People	guy/guys (2,238), kid (190)
3. Time	time/times (4,002)
4. Events	call (453), sell (280), crime (270), fight (246), competition (231), break (218), battle (165)
5. Places	spot (365), road (287)
6. Objects	stuff (219), nails (201)
7. Economy	economy (200)

Results revealed that the adjective “*tough*” shows a clear semantic preference for collocating with nouns representing abstract concepts, people, time, events, places, objects, and economy. The classification of the top 30 noun collocates from COCA into seven semantic categories was guided by dictionary definitions and concordance line analysis to capture the nuanced meanings of “*tough*” in various contexts.

The most prominent category is Abstract Concepts, which includes high-frequency collocates such as “*question/questions*”, “*decisions/decision*”, “*choices/choice*,” “*love*,” “*job*,” “*situation*,” “*issues*,” “*luck*,” “*loss*,” and “*position*”. These nouns highlight mental, emotional, or situational difficulty and complexity, emphasizing “*tough*” as expressing challenge, uncertainty, or hardship in non-physical domains. For example, phrases like “*tough question*” or “*tough decisions*” suggest cognitive or emotional strain rather than physical hardness.

The People category, including collocates such as “*guy/guys*” and “*kid*” reflects how “*tough*” often describes human qualities related to strength, resilience, or demeanor, showing its frequent use in characterizing individuals perceived as strong or unyielding.

The Time category, represented by “*time/times*” underscores the adjective’s role in describing periods marked by difficulty or adversity, as in “*tough times*” which conveys endurance through challenging temporal phases.

Event-related nouns such as “*call*,” “*sell*,” “*crime*,” “*fight*,” “*competition*,” “*break*,” and “*battle*” show that “*tough*” frequently characterizes actions or occurrences involving struggle, confrontation, or high stakes, thus reinforcing the association of the adjective with challenge and resistance in dynamic contexts.

The Places category, with nouns like “*spot*” and “*road*” typically refers to

physical locations or metaphorical points of difficulty, while the Objects category, including “*stuff*” and “*nails*” suggests physical toughness or durability but is less dominant.

Lastly, the Economy category, with “*economy*” as a collocate, points to the use of “*tough*” in describing economic conditions or environments that are difficult or demanding.

Overall, the findings demonstrate that “*tough*” predominantly collocates with nouns that signify abstract difficulty, human characteristics, challenging situations, and adversities, reflecting its semantic role in expressing endurance, strength, and complexity across various domains. There are some examples of this group of semantics relating to *tough* found by retrieving the data from concordance lines of the COCA, which are presented as follows in (3), (4), (5) and (6).

(3) *on the cost of the damage done by Sandy. # " There is a tough time, " Gillian said. " I've never seen a storm like this*

(4) *the important stuff and that they are still free to spend frivolously. It's tough love, but letting them live without electricity for a few days is a great*

(5) : *Thank you for being (unintelligible) -- (Cross-talking CRAIG-MELVIN) : It was a tough loss over the weekend for (unintelligible) Saturday.*

(6) *Tebow during the Broncos' second-half surge in 2011. # " It was a tough situation, " Elway said, " There are Tebow fans and there are Broncos*

In excerpt (3), “*tough time*” fits the Time category, describing a difficult period marked by adversity and hardship caused by the storm; in excerpt (4), “*tough love*” belongs to the Abstract Concepts category, expressing emotional difficulty and a challenging but caring attitude; excerpt (5), “*tough loss*” also falls under Abstract Concepts, indicating an emotionally painful experience; and excerpt (6), “*tough situation*” refers to a challenging circumstance, again within Abstract Concepts, emphasizing complexity and struggle in a non-physical domain.

4.2.3 Semantic Preference of “*Difficult*”

To gain a deeper understanding of the usage of the adjective “*difficult*”, the top 30 noun collocates from Table 4.1.3 were grouped into semantic categories based on their shared meanings and contextual patterns. The classification process involved

consulting dictionary definitions and examining concordance lines. As a result, the collocates were organized into Five (5) main semantic categories:

Category	Nouns
1. Abstract Concepts	situation/situations (1,570), decision/decisions (1,374), task/tasks (1,363), questions (607), position (609), issues (462), choices (338), circumstances (466), conditions (292), proposition (85), aspect (73), topics (60)
2. Time	times (1,301), period (328), moments (125)
3. Events	challenge/challenges (509), transition (192), journey (155), conversations (104), negotiations (66), balancing (60), adjustment (53)
4. Place	terrain (137)
5. Objects	gauge (139)

Results revealed that the adjective “*difficult*” shows a clear semantic preference for collocating with nouns representing abstract concepts, time, events, places, and objects. The classification of the top 30 noun collocates from COCA into five semantic categories was guided by dictionary definitions and concordance line analysis to capture the nuanced meanings of “*difficult*” in various contexts.

The most prominent category is Abstract Concepts, which includes high-frequency collocates such as “*situation/situations*”, “*decision/decisions*”, “*task/tasks*”, “*questions*”, “*position*”, “*issues*”, “*choices*,” “*circumstances*”, “*conditions*”, “*proposition*”, “*aspect*”, and “*topics*”. These nouns highlight mental, emotional, or situational complexity and challenge, emphasizing “*difficult*” as expressing cognitive effort, problem-solving, or complicated circumstances. For example, phrases like “*difficult decisions*” or “*difficult situations*” suggest the presence of complexity or obstacles that require thoughtful consideration.

The Time category, represented by collocates like “*times*”, “*period*”, and “*moments*” underscores the adjective’s role in describing specific durations or points marked by hardship or complexity, as in “*difficult times*,” conveying endurance through challenging temporal phases.

Event-related nouns such as “*challenge*”, “*transition*”, “*journey*”,

“conversations”, “negotiations,” “balancing” and “adjustment” indicate that “difficult” frequently characterizes dynamic processes or experiences involving struggle, change, or adaptation, highlighting its use in marking complexity in ongoing actions or interactions.

The Places category, with “terrain” as a collocate, refers to physical environments that are hard to navigate or endure, while the Objects category, including “gauge,” reflects physical tools or instruments that are perceived as hard to use or understand.

Overall, the findings demonstrate that “difficult” predominantly collocates with nouns that signify complexity, challenge, or effort, particularly within abstract and situational domains, underscoring its semantic role in expressing intellectual, emotional, and physical obstacles across various contexts.

Below are some examples in this group of semantics relating to “difficult” found by retrieving the data from concordance lines of the COCA, which are presented as follows in (7), (8), (9), and (10).

(7) ...gave birth to it. Although these netizens are sympathetic towards women who face the difficult **decision** to get an abortion, they do not explicitly refer to their right to...

(8) Your mother would say we are stained by such choices. Such brutal and difficult **choices** we've made, but we can only do our best.

(9) In 1985 because the Diet determined to enact the law before ratifying CEDAW. The difficult **negotiations** that involved the Ministries of Labor and Education, as well as unions.

(10) # To be fair to the Federalis, they are in a difficult **position**. The U.S. capital market regulatory structure is a hierarchical command-and-control process.

Examining excerpts (7) to (10). In excerpt (7), “difficult decision” refers to a complex and emotionally challenging choice, illustrating the Abstract Concepts category where “difficult” describes cognitive and moral complexity involved in making decisions. Excerpt (8) uses “difficult choices” to highlight emotionally taxing decisions, reinforcing the abstract and situational hardship conveyed by “difficult.” In excerpt (9), “difficult negotiations” exemplifies the Events category, where “difficult”

characterizes complex interactions requiring effort and compromise among parties. Lastly, excerpt (10) using “*difficult position*” reflects the Abstract Concepts category again, portraying a challenging circumstance or state that demands careful navigation. These examples illustrate how “*difficult*” consistently modifies nouns that denote complexity, challenge, or demanding situations, emphasizing mental, emotional, or procedural effort across various contexts.



CHAPTER 5

DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

This chapter presents the summary of findings, discussion of results, conclusion of the study, and the implications and recommendations based on the data analysis.

5.1 Summary of the Finding

Analyzing the similarities and differences between the adjectives “*hard*”, “*tough*”, and “*difficult*” with an emphasis on their collocational patterns and semantic preferences was the goal of this study. These adjectives were selected because they describe a variety of difficulties and are often used in both written and spoken English. The study's primary research tools were the Corpus of Contemporary American English (COCA) and the online dictionaries Merriam-Webster and the Longman Dictionary of Contemporary English. Definitions were supplied by the dictionaries, and comprehensive information on collocations was provided by COCA, a sizable corpus of more than a billion words from a variety of genres. Taking into account the adjective's frequency and Mutual Information (MI) score of higher than 3, the data analysis concentrated on finding and examining the top 30 noun collocates of these adjectives. Then, the adjectives “*hard*”, “*tough*”, and “*difficult*” were analyzed and classified their associated semantic fields. The findings reveal that the adjectives under study have a lot of parallels and distinctions. Although the terms “*hard*”, “*tough*”, and “*difficult*” are synonymous, not all circumstances call for the usage of all three terms.

5.1.1 Noun Collocations

Collocations show different collocates of each synonym and can give important insight into the semantic preference of each synonym. By looking at the data provided by COCA, it is possible to identify the most frequent and significant nouns that commonly occur with “*hard*”, “*tough*”, and “*difficult*”, revealing subtle differences in their usage. For “*hard*”, the most frequent collocates include “*work*”

(11,813), “time” (11,108), and “drive” (3,736), with “cafe” (199) and “lesson” (182) as the least frequent. The MI scores for “hard” range from 3.05 to 7.16, indicating a strong association with both general and specific contexts. For “tough”, frequent collocates include “time” (2,087), “guy” (1,780), and “decisions” (868), while “battle” (165) and “position” (183) occur less often. The MI scores for “tough” range from 3.01 to 7.00, showing strong associations in particular contexts. For “difficult”, high frequency collocates are “times” (1,301), “task” (1,160), and “situation” (1,151), while “concept” (77) and “proposition” (85) are the least frequent. The MI scores for “difficult” range from 3.26 to 6.20, reflecting its broad use in both general and specialized contexts.

5.1.2 Semantic Preference

Semantic preference reveals how frequently occurring collocates of a word tend to fall into certain meaning-based categories, providing insight into the conceptual associations of that word in use. For the adjectives “hard”, “tough”, and “difficult”, their top 30 noun collocates were grouped according to semantic fields, based on dictionary definitions and concordance line analysis from the COCA corpus. These semantic categories help illustrate how each synonym relates to specific types of experiences, objects, or concepts. For “hard”, collocates were classified into eight semantic categories: Abstract Concepts, Time, Objects, Economy, People, Events, Places, and Knowledge. The dominant category was Abstract Concepts, with frequent collocates like “work,” “feelings,” “choices,” and “decisions,” emphasizing “hard” as describing mental, emotional, or moral exertion. Time and Objects like “hard disk” or “hard plastic”, display both metaphorical and literal senses of difficulty and physical firmness. Other categories such as Economy “hard currency”, “hard labor” and People “hard worker”) reinforce the association with intensity, effort, or resilience.

In contrast, “tough” collocates fell into seven categories, these are Abstract Concepts, People, Time, Events, Places, Objects, and Economy. The most prominent semantic field was Abstract Concepts, including collocates like “question”, “decisions”, “love”, and “situation,” which point to emotional or situational hardship. The People category “tough guy” highlights its use in describing human resilience or strength. Time “tough times” and Events “tough fight” “tough competition” emphasize

endurance and struggle across dynamic circumstances, while Places and Objects “tough road,” “tough nails” relate to physical toughness and durability.

For “*difficult*”, five semantic categories emerged: Abstract Concepts, Time, Events, Place, and Objects. The dominant category was Abstract Concepts, including “decision”, “task”, “situation”, and “questions” illustrating “difficult” as primarily concerned with cognitive and evaluative complexity or problem-solving. The Time category “difficult times” and Events “difficult negotiations,” “difficult transition” highlight its role in describing challenging periods and dynamic processes. Places and Objects “difficult terrain”, “difficult gauge” were less frequent but indicate physical or practical challenges.

5.2 Discussion

5.2.1 Comparing Collocations of Adjectives “*Hard*”, “*Tough*”, and “*Difficult*”

The analysis of noun collocations in terms of frequency and Mutual Information (MI) values helps to identify the strength and regularity of association between each of the adjectives “*hard*”, “*tough*”, and “*difficult*”) and the nouns they modify. This section compares and contrasts the frequency and MI value patterns for each synonym based on the data from the COCA. This allows for an objective comparison of how often each adjective co-occurs with particular nouns and the strength of those associations.

The adjective “*hard*” has the highest noun collocate frequencies among the three adjectives. Its most frequent collocates are “*work*” (11,813), “*time*” (11,108), and “*drive*” (3,736). In comparison, “*tough*” shows “*time*” (2,087), “*guy*” (1,780), and “*decisions*” (868) as its most frequent collocates, while “*difficult*” lists “*times*” (1,301), “*task*” (1,160), and “*situation*” (1,151). These figures indicate that “*hard*” co-occurs with certain nouns much more frequently than “*tough*” and “*difficult*”. In terms of MI values, “*hard*” also presents high values such as “*knocks*” (7.16), “*disk*” (7.13), and “*drives*” (6.20), showing strong associations in specific contexts. Likewise, “*tough*” demonstrates high MI values with “*nails*” (7.00), “*choices*” (6.53), and “*decisions*” (6.44), while “*difficult*” shows strong collocational strength with “*gauge*” (6.20), “*terrain*” (5.65), and “*situations*” (5.48). These MI values suggest that all three adjectives form strong lexical combinations with specific nouns, though the patterns

differ in frequency.

In addition, “*hard*” includes collocates such as “*lesson*” (182) and “*cafe*” (199), while “*tough*” features “*battle*” (165) and “*position*” (183). In contrast, “*difficult*” has “*concept*” (77) and “*proposition*” (85) as its least frequent collocates. Among these, the adjective “*difficult*” shows the lowest frequencies, suggesting a narrower distribution in some collocations. The minimum MI values also vary slightly: “*hard*” shows “*hearing*” (3.05), “*tough*” includes “*call*” (3.01), and “*difficult*” presents “*issues*” (3.26) at the lower end. Despite these values being relatively close, they reflect slightly different patterns in collocational strength among less frequent nouns.

Moreover, some nouns also appear across all three adjectives. “*Time*” or “*times*” is a frequent collocate for “*hard*” (11,108), “*tough*” (2,087), and “*difficult*” (1,301). Other shared collocates include “*decisions*”, “*choices*”, “*questions*”, and “*situation*”, although their frequencies and MI values differ. For instance, “*decisions*” appears with “*hard*” (424), “*tough*” (868), and “*difficult*” (641), while “*choices*” is seen with “*hard*” (655), “*tough*” (591), and “*difficult*” (338). These shared nouns indicate some overlap in usage and suggest that the adjectives can occur in similar syntactic environments, but with varied frequency and collocational strength.

In contrast, some noun collocates are only associated with one adjective. In adjective “*hard*” includes nouns such as “*disk*”, “*worker*”, and “*cash*”, which do not appear in the lists for “*tough*” or “*difficult*”. “*Tough*” shows unique nouns such as “*guy*”, “*nails*”, and “*fight*”, while “*difficult*” presents “*gauge*”, “*terrain*”, and “*proposition*”. These noun collocates appear only with one adjective in the data and highlight the distinct collocational profiles of each synonym.

Overall, “*hard*” has the highest overall frequencies, indicating broad and frequent use. “*Tough*” displays lower frequencies but includes strong collocational ties with particular nouns, as reflected in its MI values. “*Difficult*” shows moderate frequency levels and consistent MI values, suggesting both general and specialised usage. These findings show that while the three adjectives share some collocates, they also exhibit unique patterns in noun combinations, which will be further explored in the following section on semantic preferences.

5.2.2 Comparing Semantic Preference of Adjectives “Hard”, “Tough”, and “Difficult”

Beyond identifying frequent collocates, it is also important to understand the semantic preferences these collocates reflect. To do this, we must examine the typical meaning categories associated with the adjective. According to Stubbs (2001), semantic preference refers to the tendency of a word to co-occur with lexical items that belong to particular semantic fields. Hence, this section discusses how adjectives “hard”, “tough”, and “difficult” align with specific domains of meaning, supported by dictionary definitions and concordance lines, offering a clearer understanding of how these near-synonyms function in real-world usage. From the examined top 30 noun collocates found in the COCA corpus. These nouns were classified into categories to show what kind of meanings each adjective usually connects with. Although all three adjectives describe something that is not easy, the collocates reveal differences in how they are used.

The adjective "hard" shows a wide range of meanings, with collocates grouped into eight categories: Abstract Concepts, Time, Objects, Economy, People, Events, Places, and Knowledge. The most common group was Abstract Concepts, with words like "work", "feelings", "choices", and "decisions". These show that "hard" is often used to describe mental, emotional, or moral effort. This fits the Oxford Dictionary definition of "requiring a lot of effort or endurance". In the Time category, phrases like "hard times" or "hard year" suggest long periods of difficulty. In the Objects category, terms such as "hard plastic" and "hard disk" show both the literal meaning something physically solid and the technical use like in computer parts. Other examples include Economy, with phrases like "hard currency" or "hard labor", and People, like "hard worker". In Knowledge, "hard facts" also suggests something strong or firm. All these show that "hard" is often used when talking about strength, effort, and persistence, both in physical and mental ways.

In comparison, "tough" is linked to seven categories: Abstract Concepts, People, Time, Events, Places, Objects, and Economy. Like "hard", its most common group was Abstract Concepts. Words like "question", "decisions", "love", and "situation" show that "tough" often describes emotional or situational problems, things that are painful, stressful, or morally challenging. "Tough" is also often used to describe people, such

as "tough guy" or "tough kid", which is not common with "hard" or "difficult". In Time and Events, we see phrases like "tough times", "tough fight", and "tough competition", which suggest struggle or conflict. This supports the Cambridge Dictionary definition of "tough" as "difficult to do or deal with", especially in life's hardships. In physical examples like "tough nails" or "tough road", the word also shows strength and durability.

The adjective "difficult" has a smaller but more formal use, with collocates grouped into five categories: Abstract Concepts, Time, Events, Places, and Objects. The most common category was Abstract Concepts, including "decision", "task", "situation", and "question". This matches the Merriam-Webster Dictionary definition of "difficult" as "hard to do, deal with, or understand", showing that it often refers to things that require thinking, learning, or problem-solving. The Time category such as "difficult times" and Events like "difficult negotiations" or "difficult transition" show its use in describing complex or uncomfortable changes. Although "difficult" sometimes appears in Places and Objects, like "difficult terrain" or "difficult gauge", these are less frequent, and the word is not usually linked to physical strength or endurance.

Interestingly, there are also categories that all three adjectives share. For example, Abstract Concepts and Time appear in all three, especially when talking about decision-making or emotional situations. Words like "choices", "decisions", and "situation" can be used with all three adjectives: "a hard decision", "a tough decision", or "a difficult decision". These phrases may seem similar, but there are small differences in meaning, "hard" suggests effort, "tough" implies emotional stress or conflict, and "difficult" points to mental or complex challenges.

At the same time, each adjective also has its own special use. "Hard" is the only one often used for physical things like "rock", "plastic", or "drive", and in technical terms like "hard disk". "Tough" is unique in describing people and personality, like "tough guy", "tough kid", or "tough negotiator". Conversely, "difficult" is typically employed with collocates like "concepts," "negotiations," "propositions," and "adjustments" in formal, intellectual, or mental settings.

5.3 Conclusion

The following conclusions below are drawn from the results and discussions.

1. The findings indicate that “*hard*”, “*tough*”, and “*difficult*” are synonyms that generally describe challenges or difficulties. However, these adjectives cannot be used interchangeably in all contexts. Each adjective has its own differences in meaning and usage, which makes it crucial to select the most appropriate word based on the context, the type of challenge, and the specific semantic field it aligns with.

2. The analysis of noun collocations shows that “*hard*” has the highest frequency of noun pairings, indicating broad usage, especially with abstract concept related terms. “*Tough*” and “*difficult*” have lower frequencies but form strong associations with specific nouns, such as *guy* for “*tough*” and “*proposition*” for “*difficult*”. While all adjectives collocate to noun “*decisions*” and “*time*”, each adjective also has distinct collocational patterns that reflect their unique usage contexts.

3. Despite some overlap in shared semantic fields like Abstract Concepts and Time, each adjective also shows unique patterns. “*Hard*” appears in physical and technical terms, “*tough*” strongly connects with descriptions of people and struggle, and “*difficult*” is preferred in cognitive and formal domains.

4. This study emphasizes the importance of using corpora, in analyzing language use. Traditional dictionaries alone are not sufficient for capturing the full range of meanings and collocational patterns that words exhibit in real-world usage. Corpus data allows for a more comprehensive understanding of how words function in context, providing valuable insights for language learners, educators, and researchers.

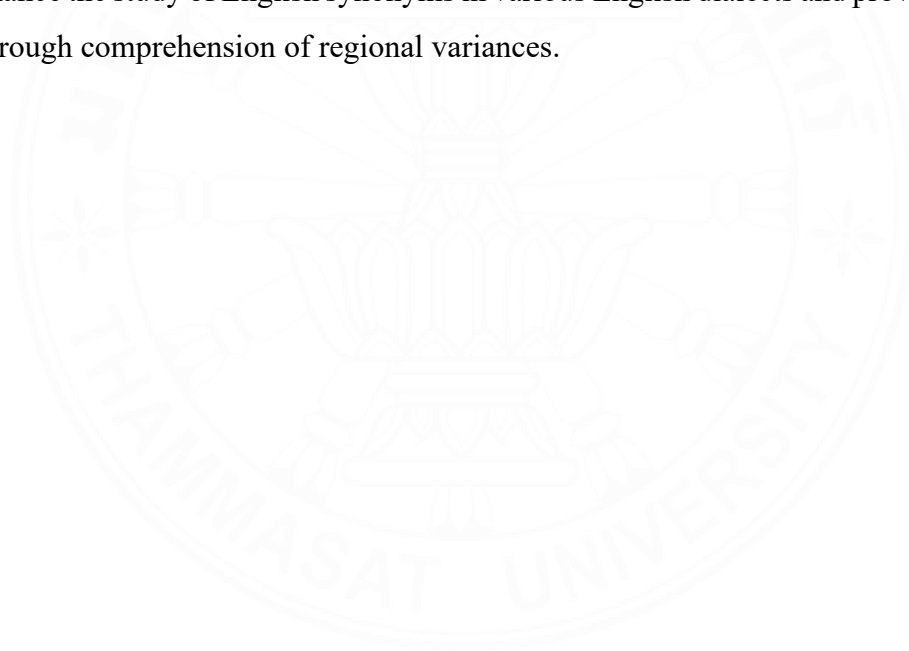
5.4 Implications

The findings of this study can help teachers and students see the parallels and discrepancies between the adjectives “*hard*”, “*tough*”, and “*difficult*”. It can also help learners use these synonyms more accurately and appropriately in various contexts, especially in writing and academic tasks. Furthermore, the study shows the important of using corpus based analysis in distinguishing near-synonyms, encouraging both teachers and students to apply similar methods in vocabulary learning. Therefore,

these findings are particularly useful in English language classrooms, especially for learners at the secondary level who are working toward mastering vocabulary use. By incorporating corpus data into teaching practices, educators can enhance students' awareness of semantic and collocational differences among synonyms.

5.5 Recommendations

This study relied primarily on data from COCA and selected dictionary sources, which are based on American English. As such, the findings may reflect usage specific to American contexts. To gain a broader perspective, future research could incorporate additional corpora such as the British National Corpus (BNC) or the Corpus of Global Web-Based English. Including or incorporating a wider range of data sources will enhance the study of English synonyms in various English dialects and provide a more thorough comprehension of regional variances.



REFERENCES

- Bennett, R. (2010). *Corpus linguistics and language teaching*. Cambridge University Press.
- Cai, J. (2012). A corpus-based study of the synonyms of “great.” *International Journal of Linguistics*, 4(2), 112–121. <https://doi.org/10.5296/ijl.v4i2.1562>
- Chalker, S., & Weiner, E. (1998). *Oxford dictionary of English grammar*. Oxford University Press.
- Chen, M. (2023). A corpus-based analysis of the synonyms "funny", "humorous", and "hilarious". *Journal of English Language Studies*, 15(1), 45–62.
- Danglli, L., & Abazaj, G. (2014). Lexical cohesion, word choice and synonymy in academic writing. *Mediterranean Journal of Social Sciences*, 5(14), 628–632. <https://files.eric.ed.gov/fulltext/EJ1348573.pdf>
- Erika, T. (2019, June). *The power of synonyms: The good, the bad, and the in-between*. University of Waterloo. <https://uwaterloo.ca/writing-and-communication-centre/blog/power-synonyms-good-bad-and-between>
- Froland, M. (2024, March 28). *Unpacking size descriptors: ‘Big’, ‘large’, ‘huge’, ‘enormous’ & ‘giant’*. Two Minute English. <https://twominenglish.com/big-large-huge-enormous-giant/>
- Harley, T. (2006). Speech errors: Psycholinguistic approach. In *Elsevier eBooks* (pp. 739–745). <https://doi.org/10.1016/b0-08-044854-2/00797-5>
- Harmer, J. (2007). *The practice of English language teaching* (4th ed.). Pearson Education. https://coljour.files.wordpress.com/2018/09/jeremy_harmer_the_practice_of_english_language_teaching_4th_edition_longman_handbooks_for_language_teachers.pdf
- Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach*. (pp. 47–69). Language Teaching.
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Routledge.
- Hornby, A. S. (2018). *Oxford Advanced Learner’s Dictionary of current English* (9th ed.). Oxford University Press.

- Hunston, S., & Francis, G. (2000). *Pattern grammar: A corpus-driven approach to the lexical grammar of English*. John Benjamins.
- Jackson, H., & Amvela, E. Z. (2000). *Words, meaning and vocabulary: An introduction to modern English lexicology*. Continuum.
- Khan, R. M. I., Radzuan, N. R. M., Shahbaz, M., & Khan, I. (2018). The role of vocabulary knowledge in speaking development of Saudi EFL learners. *Arab World English Journal*, 9(1), 406–418. https://www.researchgate.net/publication/325056045_The_Role_of_Vocabulary_Knowledge_in_Speaking_Development_of_Saudi_EFL_Learners
- Lewis, M. (2000). *Teaching collocation: Further developments in the lexical approach*. Language Teaching.
- Lindquist, H. (2009). Corpus linguistics and language pedagogy: The state of the art – and beyond. *Studies in Second Language Learning and Teaching*, 1(1), 1–23. <https://doi.org/10.14746/ssllt.2009.1.1.1>
- Longman Dictionary of Contemporary English. (2009). Pearson Education Limited.
- Longman Dictionary of Contemporary English (LDOCE). (2019). *Longman dictionary of contemporary English* (6th ed.). Pearson Education.
- McCarthy, M., O’Keeffe, A., & Walsh, S. (2010). *English collocations in use*. Cambridge University Press.
- Moon, R. (2010). Using corpora to investigate the lexicon. In M. McCarthy (Ed.), *Corpus approaches to English language teaching* (pp. 1–17). Routledge.
- Murphy, M. L. (2003). *Semantic relations and the lexicon: Antonymy, synonymy and other paradigms*. Cambridge University Press.
- O’Keeffe, A., McCarthy, M., & Carter, R. (2007). From corpus to classroom: Language use and language teaching. Cambridge University Press. https://www.researchgate.net/publication/276882865_O%27Keeffe_A_McCarthy_M_J_and_R_A_Carter_2007_From_Corpus_to_Classroom_language_use_and_language_teaching_Cambridge_Cambridge_University_Press
- Oxford Advanced Learner’s Dictionary. (2005). Oxford University Press.
- Oxford Advanced Learner’s Dictionary. (2015). *Oxford Advanced Learner’s Dictionary* (9th ed.). Oxford University Press.

- Pantong Imprapai. (2023). A corpus-based study on the near-synonyms “trip”, “journey”, and “voyage.” *Journal of English Studies*, 18(2), 89–105.
- Phoocharoensil, S. (2010a). A corpus-based study of the synonyms "ask", "beg", "plead", "request", and "appeal." *Asian EFL Journal*, 12(4), 134–150.
- Phoocharoensil, S. (2010b). A corpus-based study of English synonyms. *International Journal of Arts and Sciences*, 3(10), 227–245.
- Prinzl, M. (2010). Hans Lindquist, Corpus linguistics and the description of English (Edinburgh University Press, 2009): A detailed evaluation. *Opticon* 1826, 9(1). <https://doi.org/10.5334/opt.091009>
- Sakulratanacharoen, P. (2012). Synonymy in English: A lexicological approach. *Journal of Linguistics and Language Teaching*, 3(1), 23–33.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan. <http://dx.doi.org/10.1057/9780230293977>
- Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal*, 59, 339–341. <http://dx.doi.org/10.1093/elt/cci064>
- Sinclair, J. (Ed.). (1991). *Corpus, concordance, collocation*. Oxford University Press. <https://www.scirp.org/reference/ReferencesPapers?ReferenceID=1846456>
- Sriyang, O. (2023). Formality and collocational preferences of the adjectives "rich", "wealthy", and "affluent": A corpus-based study. *English Language Teaching*, 16(3), 56–72.
- Stubbs, M. (1996). *Text and corpus analysis*. Blackwell. <https://www.scirp.org/reference/referencespapers?referenceid=2734054>
- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Blackwell. <https://www.scirp.org/reference/referencespapers?referenceid=2734055>
- Taylor, J. R. (2002). *Cognitive grammar*. Oxford University Press.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. John Benjamins.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28, 46–65. <http://dx.doi.org/10.1093/applin/aml048>
- Wilkins, D. A. (1972). *Linguistics in language teaching*. MFT Press. <https://www.scirp.org/reference/referencespapers?referenceid=1858502>
- Yang, B. (2016). A corpus-based comparative study of “learn” and “acquire.” *English Language Teaching*, 9(1), 209–220. <http://dx.doi.org/10.5539/elt.v9n1p209>

Zeng, W. (2007). *Semantic relationships between contextual synonyms*. ERIC.
<https://files.eric.ed.gov/fulltext/ED502523.pdf>

