



**A CORPUS-BASED STUDY OF ADJECTIVE
SYNONYMS:
*FUNNY, HUMOROUS AND HILARIOUS***

BY

MENGYUN CHEN

**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 2023**

**A CORPUS-BASED STUDY OF ADJECTIVE
SYNONYMS:
*FUNNY, HUMOROUS AND HILARIOUS***

BY

MENGYUN CHEN

**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 2023**

THAMMASAT UNIVERSITY

LANGUAGE INSTITUTE

AN INDEPENDENT STUDY

BY

MENGYUN CHEN

ENTITLED

A CORPUS-BASED STUDY OF ADJECTIVE SYNONYMS:

FUNNY, HUMOROUS AND HILARIOUS

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

on June 20, 2024

Chairman

Supakorn Phoocharoensil

(Associate Professor Supakorn Phoocharoensil, Ph.D.)

Member and Advisor

Supong Tangkiengsirisin

(Associate Professor Supong Tangkiengsirisin, Ph.D.)

Director

Supakorn Phoocharoensil

(Associate Professor Supakorn Phoocharoensil, Ph.D.)

Independent Study Title	A CORPUS-BASED STUDY OF ADJECTIVE SYNONYMS: <i>FUNNY</i> , <i>HUMOROUS</i> AND <i>HILARIOUS</i>
Author	Mengyun Chen
Degree	Master of Arts
Major Field/Faculty/University	English Language Teaching Language Institute Thammasat University
Independent Study Advisor	Associate Professor Spong Tangkiengsirisin, Ph.D.
Academic Year	2023

ABSTRACT

This study examines the similarities and differences of three adjective synonyms—*funny*, *humorous* and *hilarious*—based on two criteria, which are degree of formality and collocations. The Longman Dictionary of Contemporary English was used for extracting meanings and the Corpus of Contemporary American English (COCA) for gathering information across the eight genres. It was found that all the three words tend to be used in informal contexts and therefore have a low degree of formality. As for their collocations, it is found that all three words share some common nouns and adverbs, but they also possess some collocates that are unique to themselves. Thus, they cannot be used in all circumstances in the same way.

Keywords: Corpus, COCA, Synonyms, Degree of formality, Collocation

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Associate Professor Dr. Supong Tangkiengsirisin, for giving me useful advice and suggestions in the process of completing this thesis. And I would also like to show my sincere appreciation to Associate Professor Dr. Supakorn Phoocharoensil, who not only inspired me on the topic of this thesis, but also gave me insightful comments during my defense.

I am really appreciative that the Language Institute allowed me to pursue a master's degree in the past two years. The wonderful and supportive instructors and staff have provided me with great help and inspiration.

Lastly, I also wish to thank my classmates and family for providing me warmth and encouragement.

Mengyun Chen



TABLE OF CONTENTS

	Page
ABSTRACT	(1)
ACKNOWLEDGEMENTS	(2)
LIST OF TABLES	(6)
LIST OF ABBREVIATIONS	(7)
CHAPTER 1 INTRODUCTION	1
1.1 Background of the study	1
1.2 Research objectives	3
1.3 Research questions	3
1.4 Definition of key terms	3
1.4.1 Synonyms	3
1.4.2 Corpus	3
1.4.3 COCA	4
1.4.4 Concordance lines	4
1.4.5 MI score	4
1.4.6 Genres	4
1.4.7 Collocation	4
1.4.8 Degree of formality	4
1.5 Scope of the study	5
1.6 Significance of the study	5
1.7 Organization of the study	5

	(4)
CHAPTER 2 REVIEW OF LITERATURE	7
2.1 Corpus linguistics	7
2.1.1 Types of corpora	7
2.1.1.1 General corpora	7
2.1.1.2 Specialized corpora	8
2.1.2 Corpus linguistics	8
2.1.2.1 Concordancing	9
2.1.2.2 Word frequency counts or wordlists.	9
2.1.2.3 Key word analysis.	9
2.1.2.4 Cluster analysis.	10
2.1.3 Corpus linguistics in ELT	10
2.2 Synonym	11
2.2.1 Strict synonyms	12
2.2.2 Loose synonyms	12
2.2.3 Criteria for distinguishing synonyms	13
2.2.3.1 Meaning	13
2.2.3.2 Degree of formality	14
2.2.3.3 Collocation	15
2.3 Relevant studies	15
 CHAPTER 3 RESEARCH METHODOLOGY	 21
3.1 The chosen synonymous words	21
3.2 Research instruments	21
3.2.1 Longman dictionary of contemporary English	21
3.2.2 Corpus of contemporary American English (COCA)	22
3.3 Data collection	22
3.4 Data Analysis	23

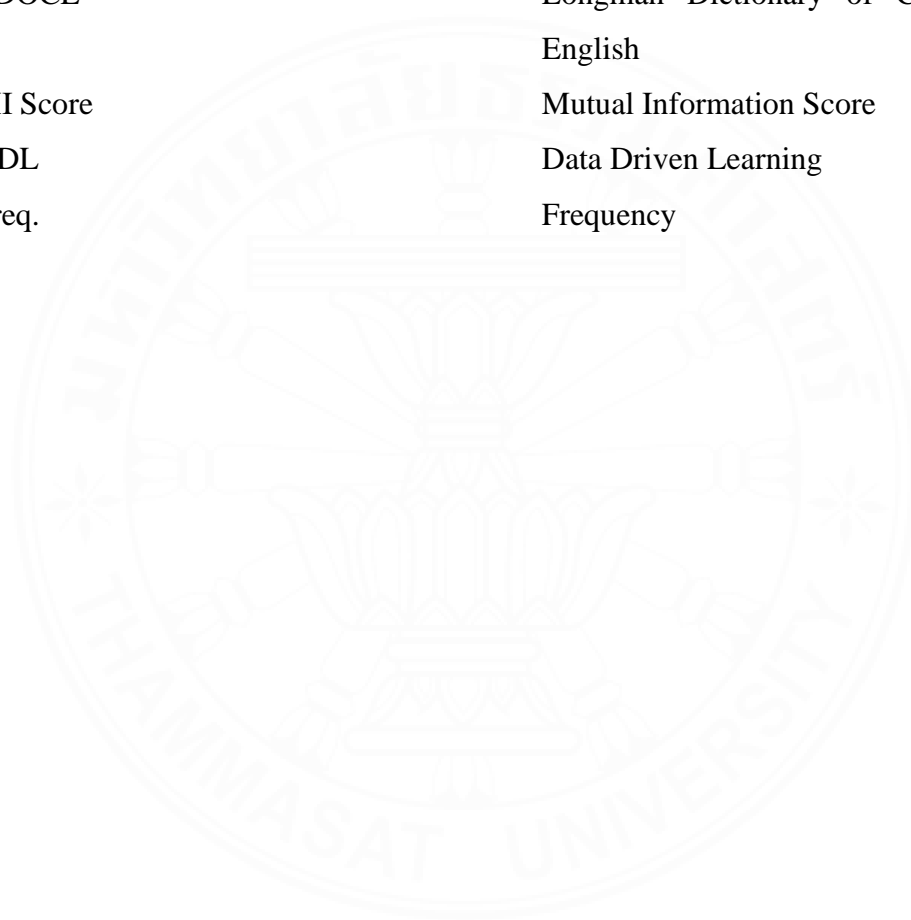
	(5)
CHAPTER 4 RESULTS	24
4.1 Meaning	24
4.2 Degree of formality	25
4.3 Collocation	27
4.3.1 Noun collocations	28
4.3.2 Adverb collocations	32
CHAPTER 5 DISCUSSION, CONCLUSION AND RECOMMENDATIONS	36
5.1 Summary of the results/findings	36
5.1.1 Degree of formality	36
5.1.2 Collocation	36
5.2 Discussion	37
5.3 Conclusion	38
5.4 Implications	39
5.5 Recommendations	39
REFERENCES	41

LIST OF TABLES

Tables	Page
4.1 Meanings and examples of adjective synonyms funny, humorous and hilarious from LDOCE	24
4.2 Overall frequency of funny, humorous and hilarious across the eight genres in COCA	25
4.3 Distribution of funny, humorous and hilarious across eight genres according to frequency	26
4.4 Top-20 frequency list of noun collocation of funny, humorous and hilarious from COCA	28
4.5 Semantic preference of noun collocates of funny from COCA	29
4.6 Semantic preference of noun collocates of humorous from COCA	30
4.7 Semantic preference of noun collocates of hilarious from COCA	31
4.8 Top-20 frequency list of adverb collocation of funny and hilarious from COCA	32
4.9 Semantic preference of adverb collocates of funny from COCA	33
4.10 Semantic preference of adverb collocates of humorous from COCA	34
4.11 Semantic preference of adverb collocates of hilarious from COCA	35

LIST OF ABBREVIATIONS

Symbols/Abbreviations	Terms
COCA	Corpus of Contemporary American English
LDOCE	Longman Dictionary of Contemporary English
MI Score	Mutual Information Score
DDL	Data Driven Learning
Freq.	Frequency



CHAPTER 1

INTRODUCTION

Seven sections make up the first chapter: the study's background, research questions and objectives, definitions of terms, scope of the study, significance of the study, and the organization of the study. The background information explains the significance of researching synonyms, the three synonym adjectives chosen for this study, and the methods employed to analyze them. The study's two main goals are outlined in the research questions and objectives. The eight specific terms found in this research are explained in the definition of terms. The scope, significance, and organization of this study comprise the remaining three sections.

1.1 Background of the Study

Given that English is the most widely used language in the world for different fields due to the development of globalization, it is considered the first global lingua franca and the language of the world. Learning English as a foreign language requires the ability in terms of the listening, speaking, reading and writing skills, and it is an undeniable fact that vocabulary forms the cornerstones of these four skills. Read (2004) pointed out that lexical items bear the fundamental information of meaning that could be understood and expressed; therefore, a limited vocabulary storage could hamper second language learners' access to the meaning of various texts and ability of effective communication in the target language. And the more vocabulary knowledge is lacking, the more it is likely to negatively impact language learners' academic performance (Jenkins et al., 1989). According to the study of Nation (2006), a vocabulary consisting of 8,000-9,000 word families is required to deal with written text, and 6,000-7,000 word families is required to deal with spoken text, assuming that 98% coverage is optimal, which reflects the vocabulary learning is a daunting process; however, the large number of synonyms makes this process even more elusive for English learners.

Synonyms are words or phrases that have similar meanings (Webb & Nation, 2017). Unlike native English speakers who have the advantage of being exposed to various meaningful contexts over time to reinforce their grasp of the subtle differences of different synonyms, students who learn English as a second or foreign language may

be limited to the synonym teaching method of new words being introduced through the ones already known, which could cause the misconception that these synonyms are interchangeable (Martin, 1984). Jackson and Amvela (2000) state that dialects, formality of the context, and connotations could be used as criteria to distinguish synonyms, and in the study of Phoocharoensil (2020), he mentioned other criteria which are genres, collocations and grammatical patterns. One way that could assist English language learners handle the similarities and differences of synonyms in terms of all these criteria is to provide them with loads of meaningful contexts, which could be achieved through the application of corpus-based data.

Sinclair (1991) compared the utility of corpora in linguistics to that of a telescope in astronomy. Thus, the fundamental justification for using corpus linguistics in language instruction is that the quality and quantity of evidence about language use made possible by corpora can—and have—led to descriptions of language that are sounder and more thorough than those made possible by earlier research techniques. Corpus linguistics has significantly contributed to the teaching of English in two major ways: the creation of reference materials for English language learners and the implementation of data-driven learning. According to Flowerdew (2013), these contributions have greatly benefited students by providing authenticity and fostering independence in their English studies. Consequently, teachers can utilize corpus-based studies of synonyms as a model to introduce corpora into the classroom, thereby encouraging students to independently learn about near-synonyms.

This study explores the similarities and differences of three synonymous adjectives: *funny*, *humorous* and *hilarious*. These three adjectives were identified as synonyms by Oxford Learner's Dictionaries and Longman Dictionary of contemporary English. The first two adjectives—funny and humorous—are included in the Oxford Dictionary's list of 3000 English words that are the most important and useful vocabulary that learners need to know, and *hilarious* is included in the Oxford's 5000 English words that are needed to be known for learners. Moreover, all three words are included in the Longman Communication 9000 words that are frequently used in English. Thus, these three adjectives are worth investigating.

Information about the target synonyms such as basic meanings was provided by Longman Dictionary of Contemporary English (LDOCE, online). Additionally, the

Corpus of Contemporary American English (COCA) was applied to collect data, such as the degree of formality and collocations that cannot be found in any of the selected dictionaries.

1.2 Research Objectives

1. To explore the similarities and differences among the synonyms *funny*, *humorous* and *hilarious* in terms of the degree of formality
2. To investigate the similarities and differences among the synonyms *funny*, *humorous* and *hilarious* in terms of common noun and adverb collocations

1.3 Research Questions

1. What are the similarities and differences among the synonyms *funny*, *humorous* and *hilarious* in terms of the degree of formality?
2. What are the similarities and differences among the synonyms *funny*, *humorous* and *hilarious* in terms of common noun and adverb collocations?

1.4 Definitions of Key Terms

1.4.1 Synonyms

Synonyms refer to the 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 *funny*, *humorous* and *hilarious*.

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). And the Corpus of Contemporary American English (COCA) is the main tool that will be used in this study to collect information about the chosen adjectives.

1.4.3 COCA

COCA (Corpus of Contemporary American English) is a database that now has over one billion words across 485,202 texts from eight genres: magazines, TV and movie subtitles, spoken language, fiction, websites, newspapers, blogs and academic journals. Through a contextual analysis, this tool assists researchers in identifying similarities and differences between synonyms.

1.4.4 Concordance Lines

Concordance lines in a corpus are lines of sentences that are often displayed with the search word or phrase—often referred to as the "node"—in the middle of the line, and seven or eight more terms on either side.

1.4.5 MI Score

MI score or mutual information (MI) value is a statistical number that is used to determine how strong a collocation is. In this study, the selection process was limited to collocates with an MI score of ≥ 3 , which denotes the important value for collocational link.

1.4.6 Genres

Genres in this study refer to eight genres that COCA has to offer: blogs, websites, TV and movie subtitles, spoken language, fiction, popular magazines, newspapers, and academic journals.

1.4.7 Collocations

Collocation refers to the nouns and adverbs that combine together with the three target adjectives: *funny*, *humorous* and *hilarious*, which frequently occur in the top-20 list and with an MI value equal to and higher than 3.

1.4.8 Degree of Formality

The degree of formality describes how formal the contexts are in which a word appears among a group of synonyms. In this study, it is based on the genres where the target synonyms appear.

1.5 Scope of the Study

This research focuses on the three adjective synonyms: *funny*, *humorous* and *hilarious*. The data were retrieved from Longman Dictionary of Contemporary English (LDOCE, online) and the Corpus of Contemporary American English (COCA), which contains the data from 1990 to 2020. This study examines the similarities and differences of the three synonyms in terms of their degree of formality and common noun and adverb collocations.

1.6 Significance of the Study

The comparison of three adjective synonyms, *funny*, *humorous* and *hilarious*, is significant for the following reasons:

- 1) This study provides a deeper understanding of the distribution, level of formality, and collocations of the three adjectives across different genres. Despite being frequent synonyms in the English language, these adjectives can be challenging to distinguish because they cannot be used interchangeably in all situations. Therefore, it is essential to increase awareness of how to recognize synonyms in English and how to appropriately use these three words in context.
- 2) This study could raise awareness of the importance of corpus linguistics and how corpora can be applied in learning and teaching synonyms. With the help of COCA, which provides a large number of authentic contexts, synonyms can be distinguished more precisely based on different criteria.

1.7 Organization of the Study

The study consists of five chapters as follows: Chapter one includes the introduction, which involves the background, research questions, research objectives, definition of terms, scope of the study, significance of the study, and organization of the study; Chapter two presents the review of literature, i.e., corpus linguistics, theories of synonyms, and related past studies; Chapter three describes the research methodology of the study; Chapter four provides the results of the data analysis; Chapter five contains a summary of findings, the discussion of the results, the conclusion, the limitations of the study, the recommendations for further research, and the pedagogical implications.

CHAPTER 2

REVIEW OF LITERATURE

This chapter is composed of three major parts: corpus linguistics, synonyms, and relevant studies. The first part discusses the two main types of corpora, which are general and specialized corpora, some basic techniques of corpus linguistics, and the application of corpora in ELT. The second section discusses two major categories of synonyms and several criteria for recognizing synonyms relevant to this study. The third part presents some examples of the relevant past studies.

2.1 Corpus Linguistics

This part discusses the two major types of corpora, which are general corpora and specialized corpora and corpus linguistics, and their application in ELT.

2.1.1 Types of Corpora

A corpus is a collection of spoken and written materials that have been electronically stored on a computer (O'Keeffe et al., 2007), and the building of corpora in the past decades was interwoven with the development of high-speed computer technology and enormous storage capacity that can process hundreds of millions of words. Linguists have amassed a wide range of corpora over time for a variety of uses, and the quantity of corpora is increasing quickly. Corpora can be used in both quantitative and qualitative analysis (Timmis, 2015). A brief summary of the most popular corpora utilized by language researchers nowadays is discussed below: general corpora and specialized corpora.

2.1.1.1 General Corpora: According to Schmitt (2013), general corpora intend to provide a widely accessible resource for baseline or comparative studies of general linguistic features, as well as to represent language in its broadest sense. In order to be balanced, a general corpus must possess diversity, containing language samples from a variety of registers and genres (Biber, 1993). An increasing number of general corpora are now designed to be quite large. Most early general corpora included only written language; however, due to technological advancements and linguists' growing interest

in spoken language, many modern general corpora also include a spoken component. This component encompasses a wide range of speech types, from casual family conversations to formal lectures and national radio broadcasts. Some of the well-known general corpora include the Bank of English (BoE), the British National Corpus (BNC), the American National Corpus (ANC), the Corpus of Contemporary American English (COCA) and the International Corpus of English (ICE) (Lindquist 2018).

2.1.1.2 Specialized Corpora: Since general corpora seek to provide a picture of the entire language, it is of necessity to build corpora for specific research objectives. As scholars come to understand the significance of register-specific language descriptions and studies, specialized corpora, or those created with more focused research objectives, might be the most important "growing sector" for corpus linguistics. A specialized corpus typically concentrates on a certain spoken or written language variant. For example, the Helsinki Corpus of English Texts is a specialized written corpus, while the Michigan Corpus of Academic Spoken English is a specialized spoken corpus. Nevertheless, there are also corpora that are composed of both written and spoken texts; the International Corpus of Learner English (ICLE) is one of the representatives, and it is also a famous example of a so-called "learner's corpus", which is one kind of specialized corpus that is becoming more and more crucial for language instructors and comprises written or spoken language samples created by non-native speakers (Schmitt, 2013).

2.1.2 Corpus Linguistics

Corpus linguistics, as a subfield of general linguistics, has gained tremendous popularity in the past decades. The origins of corpus linguistics can be found by examining the problem of observable data and the ways in which it has been addressed over time and by various theoretical schools (Lindquist, 2018). Technological advancements have provided corpus linguists with several benefits, such as the ability to gather ever-larger language samples, significantly faster and more efficient text processing and access, and user-friendly computer resources for linguistic analysis (Schmitt, 2013). Examining patterns of language use is a primary contribution of corpus linguistics. For analyzing natural language, corpus linguistics offers an incredibly

powerful tool. It can shed light on how language use varies depending on the context, such as spoken versus written or formal versus informal interactions. This has revolutionized the study of language analysis in its real context. Moreover, pedagogy has also benefited from the application of corpus linguistics in various ways, including the development of data-driven learning (DDL), course books, grammar resource books, dictionaries, English for Specific Purposes courses, and material design (Flowerdew, 2013). To achieve all these benefits of corpus linguistics, some of its basic techniques need to be examined.

2.1.2.1 Concordancing: Finding every instance of a specific word or phrase using corpus software is known as "concordancing," which is a fundamental tool in corpus linguistics. Concordance lines are often displayed with the word or phrase being searched, which are often referred to as the "node" and displayed in the middle of the line, and seven or eight more terms on either side, which are called Key-Word-In-Context displays, or KWIC-concordances (O'Keefe et al., 2007).

2.1.2.2 Word Frequency Counts or Wordlists: Another technique that corpus can do is calculating word frequency lists, also called wordlists, for any batch of texts very quickly, which would be hard to accomplish by hand. One can obtain a rank ordering of all the terms in a corpus in order of frequency by doing a word frequency list on it. This feature makes it easier to do research across various corpora, linguistic variants, and usage settings (O'Keefe et al., 2007).

2.1.2.3 Keyword Analysis: The purpose of this function is to locate important words in one or more texts. Keywords are those whose frequency is above average in relation to some standard. Keywords are typically not the words that show up most frequently in a text or set of texts. However, they are unusually frequent in another sense. To put it another way, corpus software compares two wordlists: the text users wish to study and an existing word list serving as a benchmark corpus or reference file (O'Keefe et al., 2007). A useful tool for describing a text or genre is the keyword facility. Additionally, it can be used in fields like text retrieval, content analysis, and forensic linguistics.

2.1.2.4 Cluster Analysis: The study of cluster analysis looks at how language frequently appears in groups of words or phrases (like "I mean," "I don't know," "a lot of," "a couple of," "you have to," etc.) that together describe a language's vocabulary. The process of creating chunks or cluster lists using corpus software is similar to the process of creating lists of individual words. In addition to frequency ranking, corpus software can be used to search for word combinations (O'Keefe et al., 2007).

2.1.3 Corpus Linguistics in ELT

Corpus linguistics have been utilized to inform language learning and teaching in the past decades, both indirectly and directly. Reference publishing, syllabus design and materials development, language testing and teacher development are the areas that have been influenced by corpus linguistics indirectly (Hinkel, 2011). One of the areas that have been affected by corpus linguistics enormously is the application of learner corpora. A computerized textual database of the language used by foreign language learners is called a learner corpus, and its data can be obtained from monolingual speakers or multilingual speakers. Although learner corpora theoretically might contain information obtained from participants in different contexts, in reality, many learner corpora are based on information obtained from classroom situations, such as assignments for essays or class discussions. Some of the well-known learner corpora are ICLE (International Corpus of Learner English), English Profile, and LINDSEI (Louvain International Database of Spoken English Interlanguage). Vocabulary data from learner corpora can be used in designing syllabus and graded reading materials, and to generate word lists that could help learners access the words in the list alphabetically and access words at a specific CEFR level. Learner corpora could also be useful in identifying the words that pose a challenge to learners and hence need further attention (Granger, 2002). And since a balance between frequency, difficulty, and instructional relevance must be achieved, learner corpus research can be used in this situation to assist assess how important each of these is. Moreover, test designers, examiners, and materials authors may also potentially benefit from learner corpus data. Learner corpus research could lead directly to enhanced publication outcomes in which various grammatical features are calibrated more closely with textbook preparation at a specific level, as opposed to just lexical features, as in the past (Hawkins & Buttery,

2010). Finally, learner corpora can also be applied directly to classroom instruction. Seidlhofer (2002) argues that using a "local learner corpus" of students' writings can aid in learning by addressing issues about their own or classmates' writings, as well as analyzing and fixing errors.

The usage of corpora in ELT directly is referred to as data driven learning (DDL). DDL aims to make students into linguistic researchers. Linguists have referred to learners in DDL as "language detectives" (Flowerdew, 2013). With DDL, students can work directly with corpus data, which can originate from learner corpora, specially designed teaching corpora, or general corpora such as the BNC or COCA. The teacher may supply the corpus data for DDL, or the student may obtain it on their own, under the teacher's supervision. In some circumstances, the teacher may filter the corpus data before classroom use, usually by picking the data that he or she believes is most appropriate for learners rather than changing the corpus data itself.

2.2 Synonyms

Synonymous pairs are particularly abundant in the English language. The main cause of this is the language's historical background, particularly with regard to the extensive borrowing from other languages, particularly Latin and French. The word "synonymy" is derived from the Greek word "sunonumon," which means "having the same name." It refers to a relationship of "sameness of meaning" that may exist between two words in contemporary semantics (Jackson & Amvela, 2000). Synonyms are defined in the online Cambridge Dictionary as "words or phrases that have the same or nearly the same meaning as another word or phrase in the same language, for example, 'small' and 'little' are synonyms." Nation (2002) defines synonyms as words that share similar meaning but are not always the same. For instance, the words *old*, *ancient*, *antique*, *aged*, and *elderly* are synonymous because they all denote something that is neither young nor new. But that's where the similarities end, because rather than *an ancient* or *aged record player*, we are more likely to talk about *an old record player* or even *an antique one*. Certain words are chosen over others in specific circumstances and by specific speakers, even if they appear interchangeable, like *cab* and *taxi* or *eggplant* and *aubergine*.

Jackson and Amvela (2000) categorized synonyms into two major types: strict or absolute and loose or near synonyms.

2.2.1 Strict Synonyms

Strict synonyms are interchangeable in all feasible circumstances; that is to say, substituting strict synonyms does not alter the content, style, or connotation of phrases. In reality, strict synonyms are actually quite uncommon because they would create redundancy in a language that permitted them. The English language's history indicates that when two words were extremely similar to one another and there was a good chance they would become exact synonyms, one word's meaning eventually altered or it became archaic. The way language functions and the slow, gradual evolution of word meanings eliminates strict synonyms. For example, the physical region above the ground where the sun, moon, stars, and clouds occur is referred to as "heaven" in native English; the word "sky" was later borrowed from Old Norse for this purpose. Before there was finally a noticeable difference, the two synonyms were in rivalry with one another. Today's English uses the term "sky" to refer to the firmament, while "heaven" refers to God's and the angels' spiritual domain. Similar to how *ghost* was originally English, it was in competition with the French word *spirit*, which was ultimately derived from Latin. *Spirit* now has a wider meaning than *ghost*, which is essentially limited to meanings of "disembodied spirit" (see also Holy Ghost, Holy Spirit) (Jackson & Amvela, 2000).

2.2.2 Loose Synonyms

Vocabulary that differ in minor or peripheral semantic qualities but have the same senses in terms of major semantic traits are called loose synonyms or near-synonyms (Cruse, 1986). For example, *found* and *discover* are a pair of loose synonyms that can be used as a substitute in this context: Lydia found/discovered the ball behind the garden shed. However, in the sentence "Maria Curie discovered radium in 1898," the word discovered could not be replaced by found, which means encounter something in some manner; and in the sentence "Franz found it easy to compose sonata," the word found is not interchangeable with discovered because to discover is to find something especially for the first time or something that has not been known before (Jackson &

Amvela, 2000). Given the characteristic of synonyms, it can be concluded that most of the synonyms in a language are loose synonyms instead of strict synonyms.

2.2.3 Criteria for Distinguishing Synonyms

The criteria for distinguishing synonyms include meanings, degree of formality, collocations, dialects, connotations, colligation, semantic preference, etc, and the following content will focus on elaborating three criteria related to the present study.

2.2.3.1 Meaning: Although it's unclear how this could be quantified or if there's a threshold beyond which the concept of synonymy loses all meaning, synonyms may overlap in meaning to varying degrees (Jackson & Amvela, 2000). It's critical to understand that many English words have more than one meaning. Usually, the most prevalent meaning sense will be learned first, then the less prevalent meaning senses (Schmitt, 2010). Nonetheless, acquiring further understanding about vocabulary words also requires mastering low-frequency definitions. When a word is defined adequately, its entire meaning is presented, setting it apart from the meanings of other words. Words meaning identifies the characteristics that set a symbol apart from others, lists the range of contexts in which it can be used, and clearly delineates the bounds of an idea. Both the basic and complex definitions of words are typically provided by learner's dictionaries. Accordingly, a number of earlier research studies have compared the essential and detailed meanings of words from various learner's dictionaries to identify words within a set of synonyms, such as Cambridge Advanced Learner's dictionary, Oxford Learner's dictionary, Longman Advanced American Dictionary, Longman Dictionary of Contemporary English, and Merriam-Webster Dictionary. Instructors may choose to teach a word based on its grammatical function, connections between meanings, potential conflicts between meanings, usefulness, and frequency when it has many meanings.

2.2.3.2 Degree of Formality: When two synonyms are employed, one may be used more casually in general or informal settings, whereas the other may be used in official or more formal settings; thus, the degree of formality could be considered as one criteria for distinguishing synonyms. Formal language is employed for important

objectives, such as official reports, business correspondence, and regulations, according to Svartvick and Leech (2003). Even though formal English is typically written, it can also be spoken, as in speeches given in public. While informal language is the first kind of language that a child who speaks native English is exposed to, and is employed in intimate correspondence and private conversation, which is much easier to grasp than formal language. Informal language is frequently used in public communication, including popular newspapers and commercials, usually in an informal or colloquial manner.

One way to identify words' degree of formality is to investigate their distribution across genres. Greek philosophers employed the term "genre," which meant "kind" or "form," to refer to several literary genres, including poetry, drama, and epics. Cambridge Advanced Learner's Dictionary defines "genre" as "a style, especially in the arts, that involves a particular set of characteristics" (noun) and as "produced according to a particular model or style" (adjective). Since each genre has unique characteristics and systematic descriptions of these characteristics provide learning objectives, genres are seen to be significant for pedagogy, particularly for second language learners who have had little to no exposure to first language situations. Since each word in the sets of synonyms may appear in a variety of texts and genres, genre is helpful in deciding the degree of formality of the words, which is an important criterion for separating synonyms. In her study, Sahaudon (2022) found that *accurate* is more closely related with written English or formal contexts, as the word is more commonly found in academic publications, blogs, and webpages; conversely, *correct* has a relatively lower degree of formality since it occurs more frequently in spoken English. Boontam and Phoocharoensil's (2022) study also noted that *widen* is usually used in less formal contexts, whereas *expand* and *broaden* are often associated with science and finance sub-genres, which have a higher degree of formality. The data on COCA are mainly collected from eight genres, namely blogs, websites, TV and movie subtitles, spoken, fiction, popular magazines, newspapers, and academic journals.

2.2.3.3 Collocations

Comparing synonyms with their potential collocates, i.e., recurring words or phrases and semantic prosodies, is another extremely helpful method of differentiating

between them. Benson (1985) pointed out that collocations are fixed, recurrent combinations of words in which each word basically retains its meaning. Firth (1935, as cited in O'Keeffe et al., 2007) argues that a word's meaning depends just as much on its collocations—that is, how it mixes with other words—as it does on its inherent meaning. For example, *bark* and *dog* occur so frequently together in texts, *bark* contributes to part of the meaning of *dog*, and vice-versa. Collocations frequently occur together as probabilistic events that are neither absolute nor deterministic, arising from repeated combinations used and encountered by speakers of any language. According to Thornbury (2002), even the simplest alterations to a collocation can render the text non-standard English. Collocations are therefore undoubtedly necessary to develop native-like fluency and ease of production in language. Understanding the collocational components is crucial for learners—native and second language alike. One of the advantages of corpus data is that it could demonstrate the strong statistical preference of words, which makes the repeated combinations of words to be more easily recognized.

2.3 Relevant Studies

A number of earlier research studies have attempted to evaluate and discriminate between synonyms. Many of them combine corpus data, i.e., the Corpus of Contemporary American English (COCA) and The British National Corpus (BNC), and dictionary data, e.g., *Longman Dictionary of Contemporary English*, *Cambridge Advanced Learner's Dictionary* etc, in order to perform a comparison and contrast analysis in terms of different criteria for distinguishing verb, noun, adjective and adverb synonyms.

2.3.1 A Corpus-based study of Verb English Synonyms

Using information taken from the Longman Dictionary of Contemporary English (LDOCE), Phitayakorn (2016) compared and contrasted the grammatical patterns and collocation usage of the three verb synonyms *advise*, *recommend*, and *suggest* in British English. This study found that corpus data provides greater usage information and more detailed examples than dictionary data. Furthermore, it offers a greater variety of grammatical patterns than the dictionary does. It was also

demonstrated that these three target synonyms could be used interchangeably in certain contexts but not all contexts, and they share four same grammatical patterns and nine different grammatical patterns. The study also showed that not all potential patterns of words are provided by the dictionary; therefore, concordance lines are suggested to be used as a secondary source for second language learners to offer additional patterns that are not provided in the dictionary so as to learn how to distinguish and employ these synonymous verbs more organically.

Phoocharoensil and Kanokpermpoon's (2021) study was conducted for the purpose of analyzing the genres and collocations of two high-frequency synonymous verbs *increase* and *rise* by using the data from COCA. To answer their research questions, the frequencies and distribution patterns of the two synonyms across all the eight different genres available on COCA were searched for. Additionally, they searched for nouns and adverbs with a standard of MI (mutual information) value ≥ 3 that often associated with the sample synonyms, which is a score supplied by COCA and the statistical method used in this study's collocational strength assessment. It determines if two words co-occur randomly or have a strong collocation link. If the MI value is equal to or more than 3, then the collocates would be considered as common collocates. It is found in this study that the target synonyms are linked to formal genres given their great frequency in academic works. However, because the verb *rise* is used most frequently in popular publications, it might be considered a little less formal than *increase*, which could be attributed to the reason that *rise* is more likely used to describe current trends in the vocabulary used by magazine columnists. Apart from some different adverb collocations that combine with the two target words, they also share some common adverb collocations, such as "dramatically" and "substantially." An analysis of the semantic preferences of "increase" revealed five major themes, mostly related to high speed and/or amount, which overlapped with "rise."

2.3.2 A Corpus-based study of English Noun Synonyms

Lapangdee (2017)'s study looked at potential noun collocates between the two noun synonyms "student" and "learner," as well as any discrepancies in noun-noun collocations between the two words. There were 300 concordance lines of the two target synonyms that were randomly searched for in COCA, the primary data source.

According to the results, *learner* and *student* have different collocational patterns but similar meanings overall. Additionally, it was found that the former had fewer preferred collocations than the latter. Furthermore, this study suggested that corpora offers substantial benefits for vocabulary study since they offer numerous examples of a substantial amount of the vocabulary needed to learn a language.

Phoocharoensil (2020) investigated the three noun synonyms, namely *result*, *consequence* and *outcome*, with the purpose of comparing their distribution across different genres and their common collocation drawn from COCA. The researcher selected common verb and adjective collocations that went along with the three synonymous nouns and the collocational strength as determined by statistical corpus-based approaches was used to choose the usual collocates. The mutual information (MI) value or score, is the corpus statistic used to detect collocations for this study. Collocates with an MI score of ≥ 3 , indicating significant collocational connections, were selected from the top-20 frequency list in COCA. The analysis revealed that the word "result" occurs most frequently among the three synonyms. Furthermore, because they are commonly used in academic publications, all three terms are formal in tone. However, upon closer examination of verb/adjective collocations, subtle variations in their usage preferences become apparent. It is indicated that terms with negative connotations are generally closely associated with the word *consequence*, while *result* is commonly used in research methodology and *outcome* can combine with words that have a relatively broad range of semantic properties.

2.3.3 A Corpus-based study of English Adjective Synonyms

Sahaudon's (2022) study focused on two adjective synonyms: correct and accurate. To conclude their similarities and differences in terms of their meaning, degree of formality, which are determined by their frequency and distribution across genres, collocations and semantic preference, Oxford Advanced Learner's Dictionaries and COCA were employed as data collection tools. In terms of meaning, "correct" relates to something that is true and "accurate" may refer to something that is exact; thus, they are unable to be used interchangeably in all contexts. The results revealed that the word correct was used more commonly in spoken language, which is informal contexts, while accurate was used more frequently in formal contexts, i.e., academic

texts, blogs and websites. And according to the top 30 noun and adverb collocations, selected based on a MI score of 3 or higher; although they are near synonyms, sharing three common noun collocations and four adverb collocations, they cannot always be used interchangeably or as direct substitutes for each other in all circumstances. In conclusion, it argued that dictionaries frequently fall short of giving students enough information about the distinctions and similarities between close synonyms. It's critical to understand that synonyms can differ in specific ways, even when they seem to have comparable meanings or collocations.

In Panrat's (2022) study, four adjectives—*evident*, *apparent*, *clear*, and *obvious*—were explored. These adjectives were selected due to their inclusion in the Oxford 3000, a compilation of fundamental words essential for English language learners, which features clear and obvious. Moreover, the Longman Communication 3000, encompassing the 3,000 most common words in English, includes clear, obvious, and apparent. Definitions and nuances of these adjectives were derived from the Longman Dictionary of Contemporary English online (LDOCE) and the Cambridge Advanced Learner's Dictionary (CALD, 2014). Details on collocations, formality, and grammatical patterns were supplemented by these dictionaries.

To further analyze each adjective, the Corpus of Contemporary American English (COCA) was consulted, examining their frequency and distribution across eight distinct genres. Utilizing COCA's "collocates" function, the top 30 noun and top 20 adverb collocates (with an MI score of at least 3.00) for each adjective were documented. The findings uncovered nuanced differences among these synonyms. For example, *obvious* may denote something surprising to an individual yet unnoticed by others, distinguishing it from the other three words.

In terms of formality, all four synonyms exhibit high frequency in academic writing, indicating a formal register. While they share some adverb and noun collocates, the former predominate. This study underscores the importance of corpus analysis and understanding synonym distinctions in English language education. Instructors should cultivate students' awareness of nuanced synonyms by encouraging them to consult dictionaries for various definitions and observe usage variations.

2.3.4 A Corpus-based study of Adverb English Synonyms

In his study, Techarsatitwong (2015) discussed the differences of the four linking adverbials *so*, *thus*, *therefore* and *hence* regarding their occurrences and frequency in corpus-based written academic American English, and compared and contrasted their similarities and differences between the patterns of use within the data of COCA and traditional textbooks. *Thus* had the highest frequency in academic American English, followed by *therefore*, *hence* and *so*, respectively. And there were more potential patterns found in COCA than in the restricted patterns seen in the textbooks under study. Hence, it was recommended that English language instructors integrate corpus-based data on prevalent language use patterns into their teaching, as this will enable students to experience the genuine application of these patterns and utilize the actual patterns of the linking adverbials in their own academic work.

Sormet (2017) investigated three adverb synonyms: *virtually*, *almost* and *practically*. Concordance lines were utilized to contrast definitions, collocations, and grammar patterns from COCA with those from the Oxford Advanced Learner's Dictionary (2015) and the Longman Dictionary of Contemporary English for Advanced Learners (2014). It is indicated that employing a corpus offers a nearly limitless quantity of samples for a given word to be studied. Unlike a dictionary, this text in a corpus is genuine and isn't constrained by page size or outdated. The results revealed that these three synonymous adverbs are all interchangeable and can be employed with many grammatical constructions that are not limited to using just one of the synonyms.

Based on the reviewed studies, it can be inferred that, when examining synonyms, corpus-based methods can produce more detailed findings about degree of formality, distribution patterns, semantic preference/prosody, colligation, and collocation. However, as research using a corpus to examine the similarities and differences among the synonymous adjectives *funny*, *humorous* and *hilarious* appears to be lacking, this study used corpus-based data to analyze their similarities and differences in terms of the degree of formality and noun and adverb collocations.

CHAPTER 3

RESEARCH METHODOLOGY

This part includes four sections: the target synonymous words, research instruments, data collection, and data analysis.

3.1 The Chosen Synonymous Words

The main focus of this study are three adjective synonyms: *funny*, *humorous* and *hilarious*. They were selected due to their widespread usage and high frequency in English-language communication. All three words can be found in the Longman Communication 9000, a list of the most commonly used words in spoken and written English, a collection of genuine English language corpora. Nevertheless, the first two adjectives—*funny* and *humorous*—are included in the Oxford Dictionary's list of 3000 English words that are the most important and useful vocabulary that learners need to know, and *hilarious* is included in the Oxford's 5000 English words that are needed to be known for learners. Given their shared core meanings, it may be difficult for learners to distinguish between these words. Thus, it is worthwhile to compare and analyze these three synonyms.

3.2 Research Instruments

3.2.1 Longman Dictionary of Contemporary English (LDOCE online)

The well-known American learner's dictionary, Longman Dictionary of Contemporary English, or LDOCE (online version), has 300,000 words, phrases, and definitions. It also has 88,000 pronunciation examples, one million extra corpus examples, and 165,000 corpus-based examples. One of the main characteristics of the LDOCE is the use of the 2000-word controlled defining vocabulary, the Longman Defining Vocabulary, to write all of the dictionary's definitions. Since it illustrates how written and spoken English are used differently, this dictionary was created to assist students in becoming more accurate writers and speakers of the language. Most notably, it has a feature called "Thesaurus" that aids students in growing their vocabulary by providing sets of synonyms.

3.2.2 *Corpus of Contemporary American English (COCA)*

COCA contains over a billion words in 485,202 texts, with 24-25 million words added annually between 1990 and 2019, and the texts are split equally among the following genres: blogs, websites, TV and movie subtitles, spoken language, fiction, popular magazines, newspapers, and academic journals. When studying synonyms, COCA is helpful because it offers data on the frequency of words, phrases, and grammatical constructions in various genres, which can be utilized to determine the formality level. Another special feature of COCA is its ability to search through a list of the top 60,000 words (lemmas) and examine different aspects of these words, such as synonyms, definitions, concordances, websites, clusters, and KWIC/concordance lines, enabling users to understand further about the differences among different synonyms systematically. Lastly, with COCA, users can look up the collocates of any word, which is very helpful for synonym analysis because it shows the words or phrases that each synonym in a set tends to co-occur with.

3.3 Data Collection

Firstly, the online version of Longman Dictionary of Contemporary English was used to compare the similarities and differences regarding the meaning of the three target synonyms: *funny*, *humorous* and *hilarious*.

Secondly, COCA was applied to collect concordance lines to analyze their degree of formality and noun and adverb collocations. After entering the target words in the search box and selecting the chart function, the researcher looked into the formality level as indicated by its frequency and distribution across genres. The synonyms were then placed into the search box to locate collocations and word functions before examining noun and adverb collocations. The top 20 noun and adverb collocations that most frequently occurred with the target synonyms were gathered for this study, and any collocations that had mutual information or MI scores of less than three were eliminated. This ensured that only regular, repeating collocates were retrieved, since according to Hunston (2002), for a particular collocation, the only MI scores that warrant attention are those of three or above. Then, the collected noun and adverb collocations were categorized into themes according to their semantic preference.

3.4 Data Analysis

Initially, the Longman Dictionary of Contemporary English's definitions and word usage examples of the *funny*, *humorous* and *hilarious* were examined to identify their meaning overlaps and differences. Then, COCA data on the allocation throughout eight genres was acquired in order to examine the level of formality between the three synonyms. A chart showing the overall frequency of each target word in each genre as well as the frequency of those words per million is presented based on the COCA's data. And another chart was made to present the target synonyms' distribution across the eight genres according to their relative frequency. In terms of collocations, frequency was used to represent the top 20 noun and adverb collocations. The similarities and differences among the three synonyms were then ascertained by grouping the nouns and adverbs that frequently co-occur with the synonyms into the same themes based on semantic preferences or similar meaning after the collocations had been extracted.



CHAPTER 4

RESULTS

This chapter provides an analysis of the two research questions by comparing and contrasting the three synonyms *funny*, *humorous* and *hilarious*, utilizing data from the Longman Dictionary of Contemporary English (LDOCE online version) and the Corpus of Contemporary American English (COCA).

4.1 Meaning

Table 4.1

Meanings and Examples of the Adjective Synonyms Funny, Humorous and Hilarious from LDOCE

Synonym	Meaning	Sentence example
funny	1. making you laugh	- Do you remember any funny stories about work?
	2. unusual, strange, or difficult to explain	- What's that funny smell?
	3. appear to be illegal, dishonest or wrong	- Remember, Marvin, no funny business going on while we are out.
	4. feeling slightly ill	- I always feel funny after a long car ride.
humorous	funny and enjoyable	- The film has some mildly humorous moments.
hilarious	Extremely funny	- The film was absolutely hilarious.

As presented in Table 4.1, extracted from the online version of the Longman Dictionary of Contemporary English (LDOCE), the three target adjectives share a common thread in their meanings, particularly in their ability to evoke amusement and laughter, hence indicating their status as near synonyms. However, nuances exist within their respective senses. For example, *humorous* conveys not only humor but also

suggests enjoyment, thereby imbuing a positive connotation. On the other hand, *hilarious* denotes an extreme level of amusement, setting it apart from the other two adjectives in terms of the intensity of humor conveyed. Furthermore, *funny* stands out as it encompasses meanings related to both positive and negative contexts, including situations that evoke laughter alongside those that may elicit discomfort or unease.

4.2 Degree of Formality

The degree of formality of the target synonyms *funny*, *humorous* and *hilarious* was analyzed based on their frequency in COCA in the following genres: blogs, websites, TV and movie subtitles, spoken language, fiction, popular magazines, newspapers, and academic journals. The following table shows the results of the three synonyms.

Table 4.2

Overall Frequency of Funny, Humorous and Hilarious Across the Eight Genres in COCA

Genre	Funny		Humorous		Hilarious		Total
	Freq.	Per million	Freq.	Per million	Freq.	Per million	
Blogs	10,625	82.82	644	5.01	2,101	16.34	
Webpages	9,753	78.49	907	7.30	2,095	16.86	
TV and movie subtitles	24,885	194.30	129	1.01	1,205	9.41	
Spoken	11,049	87.60	310	2.46	825	6.54	
Fiction	9,750	80.35	294	2.48	363	3.07	
Magazines	5,082	40.30	398	3.16	857	6.80	
News	4,570	37.54	518	4.25	555	4.56	
Academic texts	643	5.37	572	4.78	93	0.78	
Total	76,141		3,772		8,094		88,007

Table 4.2 presents clearly that *funny* appears most frequently in all three target adjectives, since the overall frequency of *funny* is (76,141 tokens) nine times more than that of *hilarious* (8,094 tokens) and twenty-five times more than that of *humorous* (3,772 tokens).

Table 4.3

Distribution of Funny, Humorous and Hilarious Across Eight Genres According to Frequency

Genre	Funny		Humorous			Hilarious		Total
	Freq.	Per million	Genre	Freq.	Per million	Genre	Freq.	
TV/movie subtitles	24,885	194.30	Webpages	907	7.30	Blogs	2,101	16.34
Spoken	11,049	87.60	Blogs	644	5.01	Webpages	2,095	16.86
Blogs	10,625	82.82	Academic texts	572	4.78	TV/movie subtitles	1,205	9.41
Webpages	9,753	78.49	News	518	4.25	Magazines	857	6.80
Fiction	9,750	80.35	Magazines	398	3.16	Spoken	825	6.54
Magazines	5,082	40.30	Spoken	310	2.46	News	555	4.56
News	4,570	37.54	Fiction	294	2.48	Fiction	363	3.07
Academic texts	643	5.37	TV/movie subtitles	129	1.01	Academic texts	93	0.78
Total	76,141			3,772			8,094	88,007

In Table 4.3, the distribution of occurrences for *funny*, *humorous*, and *hilarious* across various genres reveals insights into their usage contexts. *Funny* predominantly appears in TV/movie subtitles (24,885 tokens), spoken English (11,049 tokens), and blogs (10,625 tokens). Conversely, *humorous* is more prevalent in webpages (907 tokens), blogs (644 tokens), and academic texts (572 tokens), while *hilarious* is most commonly found in blogs (2,101 tokens), webpages (2,095 tokens), and TV/movie subtitles (1,205 tokens). Notably, *humorous* achieves a relatively higher rank in academic texts compared to its positions in the distribution for *funny* and *hilarious*. This suggests a greater degree of formality associated with *humorous* usage compared to the other two adjectives. Conversely, the prominent presence of *funny*, *humorous*, and *hilarious* in genres such as TV/movie subtitles, spoken English, blogs, and webpages indicates their frequent utilization in informal contexts. Considering Rungsidhaspong's (2021) classification of language in webpages, blogs, and magazines as both formal and informal, it can be inferred that these genres primarily align with informal contexts, reflecting the relatively low formality levels associated with the usage of *funny*, *humorous*, and *hilarious*.

4.3 Collocations

In this section, the top 20 noun and adverb collocations that most frequently occurred with the target synonyms were gathered and any collocations that had mutual information or MI scores of less than three were eliminated, which is to guarantee that only frequent, recurring collocates would be extracted. Then, these collocations were grouped according to their semantic meaning.

4.3.1 Noun Collocations

Table 4.4

Top-20 Frequency List of Noun Collocation of Funny, Humorous and Hilarious from COCA

Rank	Funny			Humorous			Hilarious		
	Noun collocates	Freq.	MI value	Noun collocates	Freq.	MI value	Noun collocates	Freq.	MI value
1	joke	728	4.55	story	134	3.35	joke	81	4.69
2	comedy	159	3.29	look	39	3.00	scene	69	3.17
3	accent	93	3.30	anecdote	38	8.02	comedy	62	5.23
4	comedian	82	4.30	style	30	3.50	episode	39	3.57
5	anecdote	65	4.39	account	27	3.13	tale	26	3.54
6	clown	56	3.10	tone	25	4.44	humor	22	3.80
7	bunny	37	3.06	essay	25	4.80	satire	22	6.45
8	sitcom	35	3.54	cartoon	25	5.88	comic	21	4.60
9	gag	32	3.80	caption	25	6.79	comedian	20	5.57
10	satire	28	3.49	writing	22	3.41	anecdote	20	5.99
11	one-liner	27	5.66	tale	22	4.40	adventure	18	3.30
12	skit	21	4.62	collection	21	3.14	parody	18	6.34
13	sounding	19	4.16	joke	21	3.84	commentary	17	4.19
14	prank	18	3.07	commentary	20	5.52	sketch	14	4.43
15	monologue	18	3.10	manner	19	3.73	trailer	11	3.30
16	quip	15	4.71	remark	19	4.97	one-liner	11	7.67
17	fart	14	3.49	twist	19	5.66	send-up	11	9.29
18	reefer	14	5.57	novel	17	3.30	monologue	10	5.55

19	banter	10	3.48	column	17	3.94	cameo	10	6.23
20	wordplay	6	4.28	fashion	16	3.63	spoof	9	7.58

As the results shown in Table 4.4, the three target adjectives have two common noun collocates, i.e. *joke* and *anecdote*. To compare more specifically, the noun collocates that are frequently used together with *funny* and *hilarious* are *joke*, *comedy*, *anecdote*, *satire*, *one-liner* and *monologue*, while for *funny* and *humorous*, the shared noun collocates are fewer, i.e. *joke* and *anecdote*. As for *humorous* and *hilarious*, they have *anecdote*, *tale*, *commentary* and *joke* in common. These results provide further evidence to the finding that *funny*, *humorous* and *hilarious* are near synonyms.

Table 4.5

Semantic Preference of Noun Collocates of Funny from COCA

1. HUMOR	<i>comedy, comedian, clown, sitcom, gag, skit, satire, joke, one-liner, quip, prank, banter, wordplay</i>
2. STORYTELLING	<i>anecdote, monologue</i>
3. SOUND	<i>accent, fart, sounding</i>
4. MISCELLANEOUS	<i>bunny, reefer</i>

As shown in the Table 4.5, the noun collocates of *funny* could be classified into four categories according to their semantic preference, in which the first theme HUMOR contains most of the nouns, i.e. *comedy*, *comedian*, *clown*, *sitcom*, *gag*, *skit*, *satire*, *joke*, *one-liner*, *quip*, *prank*, *banter* and *wordplay*. The second theme STORYTELLING includes *anecdote* and *monologue*. The third theme SOUND is characteristic of *funny*, containing *accent*, *fart* and *sounding*. And the two words *bunny* and *reefer* that cannot fit in the above three themes are classified as MISCELLANEOUS. The following are some example sentences extracted from COCA that can represent each of the four categories.

- 1) Of Good Stock' is a very **funny comedy** built on a tragic situation.

- 2) His family exchanged tender and **funny anecdotes** Wednesday about Goebel.
- 3) ... serendipitous is more of a **funny sounding** word to me.
- 4) ...do exactly what she wants me to do, rape, a **funny bunny** dance, or whatever.

Table 4.6*Semantic Preference of Noun Collocates of Humorous from COCA*

1. STORYTELLING	<i>account, anecdote, cartoon, caption, column, commentary, monologue, novel, remark, story, style, tale, tone, writing</i>
2. LOOK	<i>fashion, look, manner</i>
3. HUMOR	<i>joke</i>
4. MISCELLANEOUS	<i>twist, collection</i>

Table 4.6 presents the classification of noun collocates of humorous—there are two themes that are common with that of *funny*, i.e. STORYTELLING and HUMOR. However, the majority of the noun collocates are included in the theme STORYTELLING, i.e. *account, anecdote, caption, column, commentary, monologue, novel, remark, story, style, tale, tone and writing*. The second theme LOOK is specifically associated with the adjective *humorous*, including *fashion, look* and *manner*. And it is worth noting that the theme HUMOR only has one word *joke*, which is much less than the number of words under the same theme as in the situation of *funny* and *hilarious*. The fourth theme MISCELLANEOUS contains *twist* and *collection*. The following are some example sentences extracted from COCA that can represent each of the five categories.

- 1) Bosch has a wonderful, irreverent and **humorous style** to his writing that I find greatly appealing.
- 2) Poet Tony Hoagland takes a **humorous look** at the animal kingdom for some Valentine's Day verse.

3) ..., there would be some twee little ninny making coy, so-called **humorous jokes** about barcode scanners.

4) In the bedroom, a stalwart four-poster bed frame gets a **humorous twist** from pinecone finials.

Table 4.7

Semantic Preference of Noun Collocates of hilarious from COCA

1. HUMOR	<i>comedy, humor, comic, comedian, satire, parody, send-up, spoof, joke, one-liner</i>
2. STORYTELLING	<i>tale, anecdote, adventure, sketch, monologue, commentary</i>
3. TV	<i>cameo, episode, trailer, scene</i>

Table 4.7 illustrates the three themes of noun collocates of *hilarious*, and two of them are common with the themes *funny* and *humorous*, namely HUMOR (including *comedy, humor, comic, comedian, satire, parody, send-up, spoof, joke* and *one-liner*) and STORYTELLING (including *tale, anecdote, adventure, sketch, monologue* and *commentary*). And TV is a shared theme with *funny*, but it contains different words as in the situation of *funny*, i.e. *cameo, episode, trailer* and *scene*. The following are some example sentences extracted from COCA that can represent each of the three categories.

1) Maybe you should be supportive of my **hilarious jokes**.

2) She delivers **hilarious tales** about her trips to Middle East war zones, ministering to supermodels...

3) There's a **hilarious episode** in which the visceral realists attempt to kidnap Paz.

4.3.2 Adverb Collocations

Table 4.8

Top-20 Frequency List of Adverb Collocation of Funny and Hilarious from COCA

Rank	Funny			Humorous			Hilarious		
	Adverb collocate	Freq.	MI value	Adverb collocate	Freq.	MI value	Adverb collocate	Freq.	MI value
1	pretty	1090	3.20	often	146	4.02	absolutely	114	4.15
2	kind	892	3.88	sometimes	61	3.31	fucking	63	5.05
3	kinda	201	4.40	darkly	40	9.67	unintentionally	36	8.40
4	hysterically	88	7.14	somewhat	27	4.46	utterly	19	4.68
5	genuinely	63	3.43	slightly	19	3.50	downright	16	6.16
6	wickedly	57	7.35	wryly	13	8.75	darkly	11	6.70
7	hilariously	41	6.49	gently	11	4.11	alternately	8	5.76
8	darkly	39	5.23	occasionally	10	3.65	genuinely	7	3.57
9	unintentionally	34	5.01	unintentionally	9	7.50	painfully	7	4.50
10	brilliantly	30	3.90	alternately	7	6.66	wonderfully	6	4.45
11	wonderfully	28	3.37	slyly	7	8.45	outrageously	5	6.66
12	downright	27	3.61	surprisingly	6	3.47	charmingly	5	7.14
13	mildly	23	3.08	mildly	6	5.55	profoundly	4	3.28
14	outrageously	22	5.49	simultaneously	5	3.31	strangely	4	3.30
15	ridiculously	19	3.34	supposedly	4	3.20	unexpectedly	4	3.62
16	damned	16	4.62	wonderfully	4	4.97	amazingly	4	3.66
17	riotously	16	3.43	deliberately	3	3.03	delightfully	4	6.21
18	absurdly	14	8.39	wildly	3	3.56	inadvertently	4	4.09
19	insanely	13	4.05	oddly	3	3.84	devastatingly	4	7.53

Table 4.8 shows the result that the three target adjectives have three common adverb collocates, namely *darkly*, *unintentionally* and *wonderfully*. To be more specific, the adverb collocates that co-occur with *funny* and *hilarious* are *genuinely*, *darkly*, *unintentionally*, *wonderfully*, *downright* and *outrageously*, while for *funny* and *humorous*, they share fewer adverb collocates, i.e. *mildly*, *darkly*, *unintentionally* and *wonderfully*. And for *humorous* and *hilarious*, they have *alternately*, *darkly*, *unintentionally* and *wonderfully* in common.

Table 4.9

Semantic Preference of Adverb Collocates of Funny from COCA

1. DEGREE/LEVEL	<i>pretty, kind, kinda, mildly, downright</i>
2. POSITIVE FEELING	<i>genuinely, brilliantly, wonderfully, hilariously</i>
3. NEGATIVE FEELING	<i>hysterically, wickedly, darkly, outrageously, ridiculously, damned, riotously, absurdly, insanely, uproariously</i>
4. NEUTRAL FEELING	<i>unintentionally</i>

As shown in the Table 4.9, the adverb collocates of *funny* could be classified into four categories according to their semantic preference: DEGREE/LEVEL, POSITIVE FEELING, NEGATIVE FEELING, MUTUAL FEELING. And most of the adverb collocates are under the theme of NEGATIVE FEELING, including *hysterically*, *wickedly*, *darkly*, *outrageously*, *ridiculously*, *damned*, *riotously*, *absurdly*, *insanely* and *uproariously*. The following are some example sentences extracted from COCA.

- 1) I have to admit, Frank, it was **kinda funny**.
- 2) They would only laugh if it was **genuinely funny**.
- 3) I think they are **ridiculously funny**, especially the dogs that just sit and stare while looking ridiculous.

- 4) Bad movies can be pretty entertaining because of their ability to be **unintentionally funny**.

Table 4.10

Semantic Preference of Adverb Collocates of Humorous from COCA

1. FREQUENCY	<i>often, occasionally, sometimes</i>
2. DEGREE/LEVEL	<i>gently, mildly, slightly, somewhat, vaguely, wildly</i>
3. POSITIVE FEELING	<i>wonderfully, surprisingly</i>
4. NEGATIVE FEELING	<i>darkly, oddly, slyly, wryly</i>
5. NEUTRAL FEELING	<i>unintentionally, deliberately</i>
6. MISCELLANEOUS	<i>alternately, simultaneously, supposedly</i>

As illustrated in Table 4.10, the adverb collocates of *humorous* are separated quite evenly into six themes. Except for the five themes that are common with *funny*, *humorous* has one characteristic theme which is *frequency*. Words that are contained in this theme: *often, occasionally* and *sometimes*, typically collaborate with *humorous* as opposed to *funny* and *hilarious*. The following are some example sentences extracted from COCA.

- 1) Eventually, Lindelof says, Coon's distinct and **often humorous** approach changed his entire outlook on the show.
- 2) They have exaggerated, **gently humorous** features and postures.
- 3) He writes with a wonderful and **wonderfully humorous** flair for language raising his work above most crime genre offerings and bumping it right.
- 4) ...scandals in the publishing world form the backbone of this fast-paced and **wryly humorous** mystery.
- 5) offer tiny glimpses of Clinton and her correspondents at perhaps their most unguarded and **unintentionally humorous**.
- 6) ...features **alternately humorous** and doleful tales from the brothers' fascinating lives in Ireland.

Table 4.11*Semantic Preference of Adverb Collocates of Hilarious from COCA*

1. DEGREE/LEVEL	<i>Absolutely, utterly, downright</i>
2. POSITIVE FEELING	<i>Genuinely, wonderfully, charmingly, profoundly, amazingly, delightfully, awesomely</i>
3. NEGATIVE FEELING	<i>Fucking, darkly, painfully, outrageously, strangely, devastatingly</i>
4. NEUTRAL FEELING	<i>Unintentionally, unexpectedly, inadvertently</i>
5. MISCELLANEOUS	<i>alternately</i>

Table 4.11 presents the five themes of adverb collocates of *hilarious* are all shared with *funny* and *humorous*. However, unlike the situation of *funny* where most of the words are allocated in NEGATIVE FEELING, the adverb collocates of *hilarious* are distributed quite evenly in both POSITIVE (*Genuinely, wonderfully, charmingly, profoundly, amazingly, delightfully, awesomely*) and NEGATIVE FEELING (*Fucking, darkly, painfully, outrageously, strangely, devastatingly*). The following are some example sentences extracted from COCA.

- 1) He was **absolutely hilarious** and like the goofiest, just funniest way ever.
- 2) And his radio show is **genuinely hilarious** purely from a professional standpoint pace, timing, etc.
- 3) In her surreal, **darkly hilarious** and profound novel, Millet presents a disoriented hero who...
- 4) These ladies all look kind of wacktackular in an **unintentionally hilarious** way.
- 5) During my London visit, Depp is **alternately hilarious**, sly and incoherent.

CHAPTER 5

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

5.1 Summary of the Results/Findings

This study aims to explore the following two research questions:

1. What are the similarities and differences among the synonyms *funny*, *humorous* and *hilarious* in terms of the degree of formality?
2. What are the similarities and differences among the synonyms *funny*, *humorous* and *hilarious* in terms of common noun and adverb collocations?

5.1.1 Degree of formality

Based on the data in COCA, among the three target synonyms, *funny* appears most frequently in that its overall frequency is 76,141 tokens, which is nine times more than that of *hilarious*, which has 8,094 tokens, and twenty-five times more than that of *humorous*, which had only 3,772 tokens.

Funny is used in TV/movie subtitles, spoken English and blogs the most frequently, while genres that are usually deemed as formal contexts, such as magazines, news and academic texts, showed the lowest frequency. *Humorous* demonstrated the most usage in webpages, blogs and academic texts and the least usage in spoken English, fiction and TV/movie subtitles. *Hilarious* is popular in blogs, webpages and TV/movie subtitles, while news, fiction and academic texts obtained the lowest frequency.

5.1.2 Collocations

The data analysis of the top 20 noun and adverb collocates with an MI value equal to and higher than 3 from COCA have yielded further insightful information about the similarities and differences of *funny*, *humorous* and *hilarious*.

The findings indicated that there are two noun collocates shared by the three selected adjectives, which are *joke* and *anecdote*. Additionally, the collected top 20 nouns were classified into different categories regarding their semantic preference. The findings showed that these three words have some shared categories among each other,

but also have characteristic categories that specifically belong to themselves. For *funny*, there are four categories, i.e. HUMOR, STORYTELLING, SOUND and MISCELLANEOUS. It was found that nouns that are related to sound, such as *accent* and *sounding*, tend to co-occur with *funny* instead of *humorous* and *hilarious*. In terms of *humorous*, five categories have been concluded which are STORYTELLING, PEOPLE, HUMOR, TV and MISCELLANEOUS. And nouns that can be classified under the topic LOOK, such as *look*, *manner* and *fashion*, have a tendency to be employed with *humorous* instead of *funny* and *hilarious*. Regarding *hilarious*, all of its categories are shared with that of *humorous*, which are humor, storytelling and TV.

According to the statistics collected from COCA, the three adjectives have some common noun collocates: *darkly*, *unintentionally* and *wonderfully*. The adverb collocates of the three adjectives also share the same situation as their noun collocates. They have some adverb collocates in common, such as *darkly*, *unintentionally* and *wonderfully*, which are words that can be classified into different themes according to their semantic meaning. Degree/level, positive feeling, negative feeling and neutral feeling are four categories that all the three target words have in common; however, there is a category that is found specifically in the word *humorous*, which is frequency, including words such as *often*, *occasionally* and *sometimes*. Another finding is that although these three words all share the negative feeling category in terms of their adverb collocates, most of the adverbs associated with negative feelings tend to co-occur with *funny* in that the number of words assigned to this category in *funny* outnumbers that of *humorous* and *hilarious*. This correlates with the finding regarding the differences of definition of these three synonyms that apart from their shared meaning of making someone laugh, the word *funny* has additional meanings that are related to negative things and feelings.

5.2 Discussion

The results and findings imply that all the three words have a low degree of formality since TV/movie subtitles, spoken English, blogs and webpages are genres that are closely related to informal contexts and therefore can be used interchangeably in informal contexts. Nevertheless, given that the level of formality of *humorous* is found to be a little higher than *funny* and *hilarious* because of its higher ranking in

academic texts, they cannot be used interchangeably in formal contexts such as academic texts.

And according to the analysis of the collocations of the targeted adjectives, the three synonyms can be separated from one another depending on the words they most frequently occur with and they cannot be used interchangeably in all circumstances, which is consistent with Petcharat and Phoocharoensil (2017)'s study findings.

5.3 Conclusion

This research focuses on the distinction of synonyms, which is an important aspect of English language learning and instruction. The target synonyms for this study— *funny*, *humorous* and *hilarious*—are three frequently used and must-know adjectives for English learners according to some authoritative dictionaries. Among the myriad criteria for distinguishing synonyms, including dialects, connotations, and grammatical patterns, formality degree and collocations were selected as key parameters for analyzing the similarities and differences among the three synonymous adjectives. This analysis was based on data sourced from the COCA corpus, a comprehensive database comprising over one billion words drawn from 485,202 texts across eight genres. The richness and authenticity of the contexts within the COCA corpus facilitate a nuanced understanding of the varying usage nuances of these synonyms.

The results revealed that among these three synonymous adjectives, *funny* has the highest frequency across all the eight genres in COCA, followed by *hilarious* and *humorous*, respectively. And all three words have a low level of formality given the fact that they are all used mainly in informal contexts; however, the formality level of *humorous* is found to be relatively higher than *funny* and *hilarious* in that the academic texts ranks third within the eight genres.

As for the noun and adverb collocations, *funny*, *humorous* and *hilarious* share common nouns and adverbs, but each also co-occurs with some specific words, which further confirms that they are near synonyms. And even though they have some common categories that are classified according to the semantic preferences of nouns and adverbs, the noun and adverb collocates under the same categories often have a preference co-occurring with different adjectives. Therefore, with the assistance of

corpus-based data, English learners and teachers are able to identify common patterns and, more significantly, variations in their collocational behavior (Phoocharoensil, 2022).

5.4 Implications

English learners frequently turn to dictionaries when they encounter confusion regarding the usage of certain synonyms. However, dictionaries may not comprehensively capture the nuanced similarities and differences between synonyms. Consequently, students may inadvertently misuse synonyms interchangeably across different contexts. To enhance understanding of synonyms, English teachers could introduce a corpus-based approach to L2 practitioners. Leveraging corpus data offers authenticity and reliability, providing a nuanced understanding of synonyms' usage nuances. Encouraging the integration of corpus data in EFL or ESL classrooms, particularly in vocabulary instruction, is crucial. This approach fosters students' awareness of discerning synonyms in diverse contexts and with varying collocations. For instance, when teaching students how to distinguish *funny*, *humorous* and *hilarious*, teachers could start by presenting their shared meaning, which is making someone laugh, and then highlight the additional meanings of *funny* that are not found in the definitions of *humorous* and *hilarious*. This is for the purpose of making students recognize that *funny* has a broader meaning and could be employed in wider contexts. The next step could be to remind students that formality degree is also one of the important criteria in distinguishing synonyms by presenting to them the frequency of the chosen synonyms throughout all the texts and analyzing their frequency across all the genres so as to get insightful information about their degree of formality. Last but not least, collocations of the synonyms should also be emphasized to students due to the fact that near synonyms may often combine with certain words that are not shared with their counterparts in specific contexts.

5.5 Recommendations

As this study only analyzed the top 20 nouns and adverbs that often co-occur with the target synonymous adjectives, there may be other nouns and adverbs that frequently co-occur with them but were not included in the lists above due to low

frequency or MI scores in COCA. Future research could include more collocations and may discover some new findings. Moreover, only two criteria were chosen to study the chosen synonyms, future study could analyze them from a broader view, including more criteria to acquire more insightful information.



REFERENCES

- Benson, M. (1989). The collocational dictionary and the advanced learner. In M. L. Tickoo (Ed.), *Learners' Dictionaries: State of the Art* (84-91). Seameo Regional Language Centre
- Biber, D. (1993). Representativeness in corpus design. *Literary and linguistic computing*, 8(4), 243-257. https://jan.ucc.nau.edu/biber/Biber/Biber_1993.pdf
- Boontam, P., & Phoocharoensil, S. (2022). Broaden your horizons: distribution and collocational patterns of the English synonyms “Expand,” “Widen,” and “Broaden.” *The International Journal of Communication and Linguistic Studies*, 20(1), 107–123. <https://doi.org/10.18848/2327-7882/cgp/v20i01/107-123>
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge University Press.
- Flowerdew, J. (2013). *Discourse in English language education*. Routledge. <https://doi.org/10.4324/9780203080870>
- Granger, S. (2002) A bird's eye view of learner corpus research. In: Granger, S., Hung, J. and Petch-Tyson, S. (eds.) *Computer learner corpora, second language acquisition and foreign language teaching*. (pp.3-33). Benjamins.
- Hawkins, J. A., & Buttery, P. (2010). Criterial features in learner corpora: Theory and illustrations. *English Profile Journal*, 1. <https://doi.org/10.1017/s204153621000103>
- Hinkel, E. (2011). *Handbook of research in second language teaching and learning*. Routledge. <https://doi.org/10.4324/9780203836507>
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
- Jackson, H., & Amvela Zé Etienne. (2000). *Words, meaning and vocabulary: An introduction to modern English lexicology* (2nd ed.). Athenaeum Press.
- Jenkins, J. J., Matlock, B., & Slocum, T. A. (1989). Two approaches to vocabulary instruction: The teaching of individual word meanings and practice in deriving word meaning from context. *Reading Research Quarterly*, 24(2), 215. <https://doi.org/10.2307/747865>

- Lapangdee, P. (2017). *A corpus-based study of collocations of English synonyms: Student and learner*. Thammasat University.
- Lindquist, H. (2018). *Corpus Linguistics and the description of English*. Edinburgh University Press. <https://doi.org/10.1515/9781474421713>
- Martin, M. (1984). Advanced vocabulary teaching: The problem of synonyms. *The Modern Language Journal*, 68(2), 130. <https://doi.org/10.2307/327139>
- Nation, I. (1983). *Teaching and learning vocabulary*. Victoria University of Wellington English Language Institute.
- Nation, I. (2006). How large a vocabulary is needed for reading and listening? *Canadian modern language review*, 63(1), 59-82.
- Rungsidhavaspong, N. (2021). *A corpus-based analysis of English synonyms: Excellent, outstanding, superb*. Thammasat University.
- O’Keeffe, A., McCarthy, M., & Carter, R. (2007). *From Corpus to Classroom: Language use and Language teaching*. https://assets.cambridge.org/9780521616867/frontmatter/9780521616867_frontmatter.pdf
- Panrat, T. (2022). *A corpus-based study of English synonyms clear, obvious, apparent, and evident: Implications for ELT*. Thammasat University.
- Petcharat, N. & Phoocharoensil, S. (2017). A corpus-based study of English synonyms: Appropriate, proper, and suitable. *LEARN Journal: Language Education and Acquisition Research Network*, 10(2), 10-24.
- Phitayakorn, W. (2016). *A corpus-based study of English synonyms: Advise, recommend, and suggest*. Thammasat University.
- Phoocharoensil, S. (2020). A genre and collocational analysis of consequence, result, and outcome. *Journal of Language Teaching, Linguistics and Literature*, 26(3), 1–16. <https://doi.org/10.17576/31-2020-2603-01>
- Phoocharoensil, S. (2022). Primary, main, and major: Learning the synonyms through corpus data. *GEMA Online Journal of Language Studies*, 22(4). <http://doi.org/10.17576/gema-2022-2204-05>
- Phoocharoensil, S. & Kanokpermpoon, M. (2021). Distinguishing the near-synonyms ‘increase’ and ‘rise’: Genre and collocation investigation. *Kasetsart Journal of Social Sciences*, 42(4). <https://doi.org/10.34044/j.kjss.2021.42.4.34>

- Read, J. (2004). Research in teaching vocabulary. *Annual Review of Applied Linguistics*, 24. <https://doi.org/10.1017/s0267190504000078>
- Sahaudon, A. (2022). *A corpus-based study of the English synonyms 'accurate' and 'correct'*. Thammasat University.
- Schmitt, N. (2010). *Researching vocabulary*. Palgrave Macmillan
- Schmitt, N. (2013). *An introduction to applied linguistics*. Routledge. <https://doi.org/10.4324/9780203783733>
- Seidlhofer, B. (2002). Pedagogy and local learner corpora: Working with learning driven data. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 213–234). Philadelphia: John Benjamins. <https://doi.org/10.1075/llt.6.14sei>
- Sinclair, J. (1991). *Corpus, Concordance and Collocation*. Oxford University Press.
- Sormet, R. (2017). *A corpus-based study of the English synonyms almost, virtually, and practically*. Thammasat University.
- Svartvik, J., & Leech, G. (2003). *A communicative grammar of English*. Longman.
- Techarsatitwong, P. (2015). *A corpus-based study of linking adverbials of result so, thus, therefore, and hence*. Thammasat University.
- Timmis, I. (2015). *Corpus Linguistics for ELT*. Routledge. <https://doi.org/10.4324/97813157115537>
- Thornbury, S. (2002). *How to teach vocabulary*. Longman.
- Webb, S. & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.