



**A CORPUS-BASED STUDY OF THE ENGLISH
SYNONYMS ‘ACCURATE’ AND ‘CORRECT’**

BY

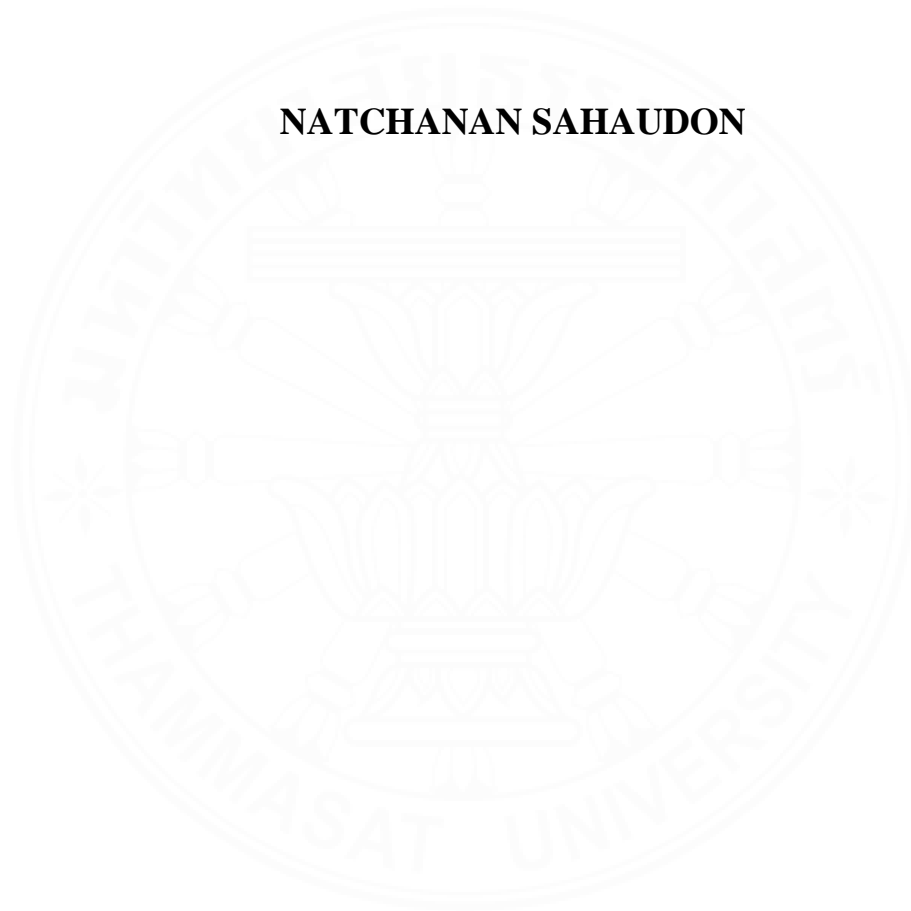
NATCHANAN SAHAUDON

**AN INDEPENDENT STUDY SUMMITTED IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF ARTS IN CAREER ENGLISH FOR
INTERNATIONAL COMMUNICATION
LANGUAGE INSTITUTE
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2022**

**A CORPUS-BASED STUDY OF THE ENGLISH
SYNONYMS ‘ACCURATE’ AND ‘CORRECT’**

BY

NATCHANAN SAHAUDON



**AN INDEPENDENT STUDY SUBMITTED IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF ARTS IN CAREER ENGLISH FOR
INTERNATIONAL COMMUNICATION
LANGUAGE INSTITUTE
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2022**

THAMMASAT UNIVERSITY

LANGUAGE INSTITUTE

AN INDEPENDENT STUDY

BY

NATCHANAN SAHAUDON

ENTITLED

A CORPUS-BASED STUDY OF THE ENGLISH
SYNONYMS 'ACCURATE' AND 'CORRECT'

was approved as partial fulfillment of the requirements for
the degree of Master of Arts in Career English for International Communication

on July 13, 2023

Chairman



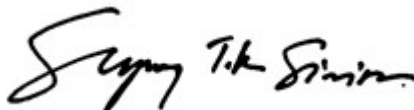
(Pimsiri Taylor, Ed.D.)

Member and advisor



(Rangsiya Chaengchenkit, Ph.D.)

Director



(Associate Professor Supong Tangkiengsirisin, Ph.D.)

Independent Study Title	A CORPUS-BASED STUDY OF THE ENGLISH SYNONYMS 'ACCURATE' AND 'CORRECT'
Author	Natchanan Sahaudon
Master of Arts	Master of Arts
Major Field/Faculty/University	Career English for International Communication Language Institute Thammasat University
Independent Study Advisor	Rangsiya Chaengchenkit, Ph.D.
Academic Year	2022

ABSTRACT

This study employed the Oxford Advanced Learner's Dictionaries and the Corpus of Contemporary American English (COCA) as data collection tools to compare and contrast the meanings, degrees of formality, collocations, and semantic preferences of the synonyms "accurate" and "correct." While these two words share a common core meaning, they exhibit differences in other aspects. In terms of formality, the COCA results, which examined the frequency of occurrences of each synonym across eight different genres, indicate that "accurate" is more frequently used in formal contexts. On the other hand, "correct" is predominantly found in spoken English, which is generally considered a more informal context. The findings on collocations and semantic preferences reveal that certain nouns and adverbs are commonly associated with both synonyms. They also share some semantic categories. However, it is important to note that even though "accurate" and "correct" are near synonyms, they cannot always be used as direct replacements or interchangeably in every situation. Although these words have similarities, their usage and appropriateness may vary depending on the degree of formality, specific collocations, and semantic contexts.

Keywords: COCA, Collocation, Corpus, Degrees of formality, Meaning, Synonyms

ACKNOWLEDGEMENTS

I appreciate my adviser, Dr. Rangsiya Chaengchenkit, for taking the time to criticize and make suggestions on how to improve my independent study. I would also like to express my gratitude to Dr. Pimsiri Taylor for her comments during the independent study defense.

In addition, I would like to thank Assistant Professor Supakorn Phoocharoensil, Ph.D., for inspiring me to conduct a corpus-based analytical research.

Finally, I want to thank my family and friends for standing by me and supporting me, enabling me to get through the challenging moments.

Miss Natchanan Sahaudon

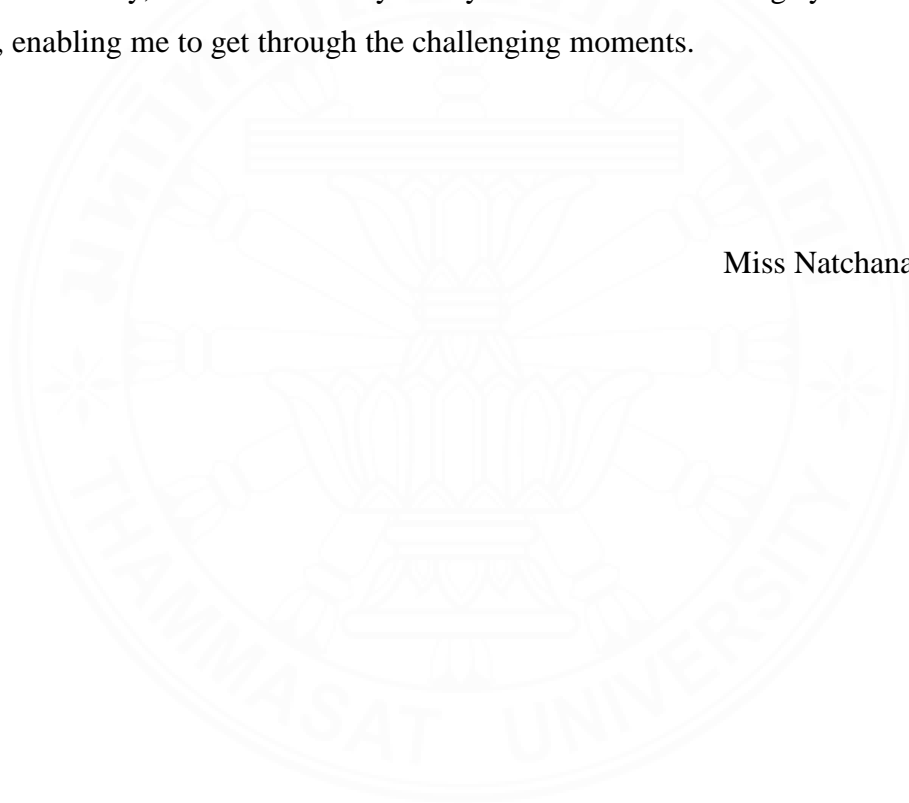


TABLE OF CONTENTS

	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	1
1.3 Research questions	2
1.4 Definition of terms	2
1.5 Significance of the study	3
CHAPTER 2 REVIEW OF LITERATURE	4
2.1 The Concept of Corpora	4
2.1.1 Definition of Corpora	4
2.1.2 The Corpus of Contemporary American English (COCA)	4
2.2 The Concept of Synonyms	5
2.2.1 Definitions	5
2.2.2 Types of synonyms	5
2.3 Criteria for distinguishing synonyms	5
2.3.1 Senses of meaning	6
2.3.2 Collocation	6
2.3.3 Semantic preference	6

	(4)
2.3.4 Formality	7
2.4 Previous related studies	7
CHAPTER 3 RESEARCH METHODOLOGY	10
3.1 The target synonymous words	10
3.2 Research instruments	10
3.2.1 Oxford Advanced Learner's Dictionary (OALD)	10
3.2.2 Corpus of Contemporary American English (COCA)	11
3.3 Data Collection and Data Analysis	11
3.3.1 Data Collection	11
3.3.2 Data Analysis	12
CHAPTER 4 RESULTS AND DISCUSSION	13
4.1 Similarities and differences in meaning	13
4.2 Degree of formality	14
4.3 Collocations	16
4.3.1 Noun Collocations	16
4.3.2 Adverb Collocations	20
CHAPTER 5 CONCLUSION AND RECOMMENDATIONS	25
5.1 Summary of the study	25
5.1.1 Objectives of the study	25
5.1.2 Selected synonyms	25
5.2 Summary of the findings	26
5.2.1 Definition of Accurate and Correct from the Oxford Advanced Learner's Dictionary	26
5.2.2 Degree of formality	26
5.2.3 Collocations	27

	(5)
5.3 Conclusion	27
5.4 Implications	28
5.5 Limitations and recommendations for future research	29
REFERENCES	30

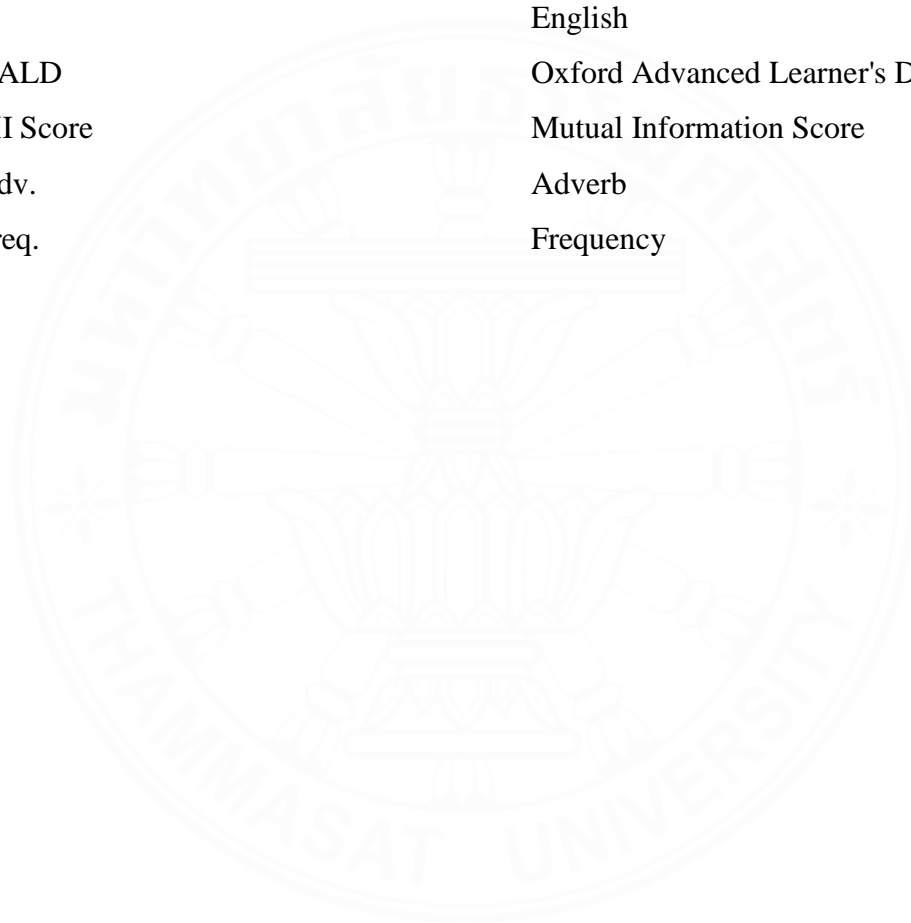


LIST OF TABLES

Tables	Page
4.1 Definition of Accurate and Correct from the Oxford Advanced Learner's Dictionary	13
4.2 Overall Frequency of Accurate and Correct across the Eight Genres from COCA	14
4.3 Distribution of Accurate and Correct across Eight Genres according to Frequency	15
4.4 Top-30 Frequency List of Noun Collocations of Accurate and Correct from COCA	16
4.5 Semantic Preference of Noun Collocates of Accurate from COCA	18
4.6 Semantic Preference of Noun Collocates of Correct from COCA	19
4.7 Top-30 Frequency List of Adverb Collocations of Accurate and Correct from COCA	20
4.8 Semantic Preference of Adverb Collocates of Accurate from COCA	22
4.9 Semantic Preference of Adverb Collocates of Correct from COCA	23

LIST OF ABBREVIATIONS

Symbols/Abbreviations	Terms
L2	Second language
COCA	Corpus of Contemporary American English
OALD	Oxford Advanced Learner's Dictionary
MI Score	Mutual Information Score
Adv.	Adverb
Freq.	Frequency



CHAPTER 1

INTRODUCTION

1.1 Background of the Study

The English language is commonly used as a global or worldwide language. One of the crucial language factors required for effective communication is vocabulary. Learners might expand their vocabulary by remembering a collection of words with a common meaning. Having a large vocabulary helps reader and listeners understand what they are reading and hearing. However, knowing the single definition of a word might not be sufficient for writing and speaking.

The use of synonyms is an essential skill that can help learners improve their English proficiency since it allows them to efficiently use a variety of words in context-appropriate ways (Rozakis, 2011). When using vocabulary in any context, learners should be aware of some words that have similar meanings. Synonyms might not be interchangeable in every situation or context, despite the fact that they share similar in meaning (Palmer, 1997). Thornbury (2002, p. 28) states that even though synonyms have similar meaning, L2 learners may become confused about how to use them in the appropriate context.

The purpose of this study is to investigate the similarities and differences the words “accurate” and “correct”, which are essential and useful words since they are frequently used in both spoken and written English and may confuse learners. As far as the researcher is aware, none of the previous studies have looked into these two synonyms. The researcher used the COCA corpus to help reveal how these two synonyms are used in different contexts.

1.2 Research Objectives

- 1) To investigate the similarities and differences between the synonyms “accurate” and “correct” in terms of meaning.
- 2) To explore the similarities and differences between the synonyms “accurate” and “correct” in terms of the degree of formality as determined by their frequency and distribution across genres.

3) To examine the similarities and differences between the synonyms “accurate” and “correct” in term of collocation and semantic preference.

1.3 Research Questions

1) What are the similarities and differences between the synonyms “accurate” and “correct” in terms of meaning?

2) What are the similarities and differences between the synonyms “accurate” and “correct” in terms of the degree of formality as determined by their frequency and distribution across genres?

3) What are the similarities and differences between the synonyms “accurate” and “correct” in terms of collocations and semantic preference?

1.4 Definitions of Key Terms

1) **Corpus (Corpora)** refers to authentic English which is used in both spoken and written communication and gathered by computer programs.

2) **COCA** stands for Corpus of Contemporary American English and refers to a database of more than one billion words from a variety of texts, such as blogs, websites, magazines, newspapers, movies, TV, spoken language, and academic texts. By examining how synonyms are used in various contexts, this tool helps researchers analyze the similarity and differences between them.

3) **Concordance lines** refer to lines of text that appear when a corpus program searches for a given word.

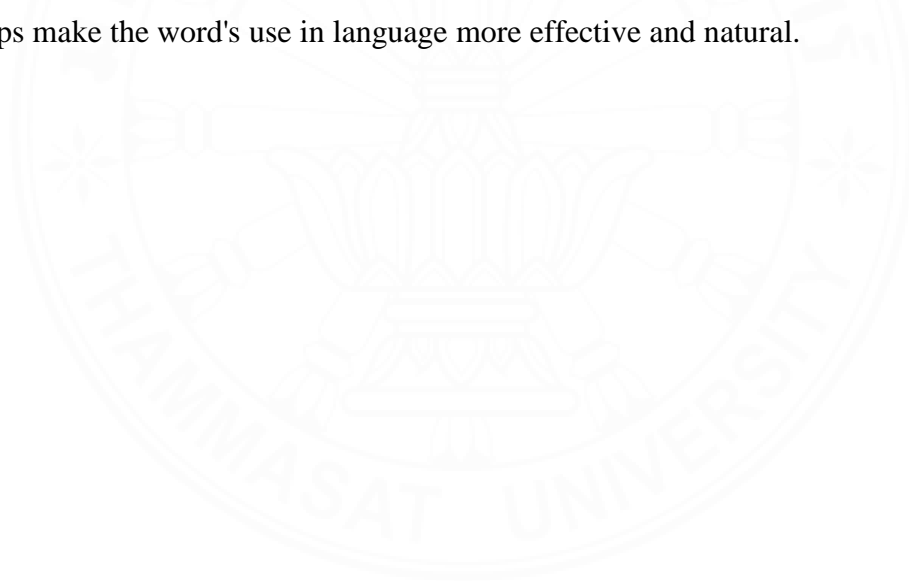
4) **Collocation** refers to words, groups of words or phrases that frequently appear together.

5) **Senses of meaning** refer to the words “accurate” and “correct”, which are synonyms, and discussed in terms of meaning in Oxford Learner’s Dictionary.

6) **Synonym** refers to the terms that in same language have an identical meaning or meanings that are similar. This study focuses on the synonyms for “accurate” and “correct”.

1.5 Significance of the Study

Vocabulary is essential for language learners. Learners use dictionaries as a tool to study vocabulary because they give the definition along with an example; although understanding the definition of a word might assist a learner understand the context, knowing the meaning alone might not be sufficient to utilize it in the appropriate context, particularly for synonyms, since the learner may become confused while deciding which one to use. Since the COCA is a highly helpful tool that can assist learners in distinguishing the similarities and differences between words or groups of words that share similar meanings, this research used it to evaluate the similarity and differences of the synonyms “accurate” and “correct”. With the use of COCA's big data on vocabulary, users can view meanings, synonyms, collocations, and distribution across different genres. Additionally, it provides real-world usage examples so that students can see the word in context and comprehend how people actually use it. This helps make the word's use in language more effective and natural.



CHAPTER 2

REVIEW OF LITERATURE

This chapter examines the relevant studies on English synonyms in order to provide answers to the three research questions on the similarities and differences between the synonyms “accurate” and “correct” in terms of meaning, formality, collocations and semantic preference. The content consists of three main sections. First, section 2.1 explores concepts and definitions of corpora, with a specific focus on COCA, or the Corpus of Contemporary American English. Section 2.2 reviews the concepts, definitions and types of synonym. As for section 2.3, it presents the criteria for distinguishing synonyms, including, sense of meaning, collocation, semantic preference and formality. Finally, the chapter ends with section 2.4, which examines the relevant empirical research.

2.1 The Concept of Corpora

2.1.1 Definitions of Corpora

Corpora is the plural form of corpus. The Oxford Advanced Learner’s Dictionary defined corpus as “a collection of written or spoken texts.” The technical definition of corpus based on the Longman Dictionary of contemporary English is “a large collection of written or spoken language, used for studying the language.” O’Keeffe, McCarthy and Carter (2007) suggest that the term “corpus” is a collection of spoken and written texts that are saved as electronic texts on computer. Today, corpus can be used to describe a sizable collection of electronic texts which is gathered from both spoken and written authentic language. We can conduct both quantitative and qualitative analysis using corpora. (Biber, Conrad & Reppen, 1998).

2.1.2 The Corpus of Contemporary American English (COCA)

The Corpus of Contemporary American English (COCA) developed by Mark Davies, a professor of linguistics, was first released in 1990. Texts totaling more than a billion words in eight genres, including spoken word, television, movies, fiction,

newspapers, academic journals, magazines, blogs and online sites are collated in COCA.

2.2 The Concept of Synonyms

2.2.1 Definitions

A synonym was described by the Oxford Advanced Learner's Dictionary as "a word or expression that has the same or nearly the same meaning as another in the same language." Synonyms refer to words or expressions that are nearly identical in meaning to other words or expressions in the same language (Thornbury, 2002). Taylor (2002) states that true synonyms are extremely rare and might not even exist.

Synonyms are useful for learners. By understanding the concept of synonyms, learners can select the appropriate words to use in the right contexts and reduce redundancy.

2.2.2 Types of Synonyms

Absolute and near synonyms are the two main types of synonyms which this study will examine.

1) Absolute synonyms refer to words that have the same meaning and are interchangeable in all contexts (Jackson & Amvela, 2000). Absolute synonyms, as defined by Cruse (2000), are two or more words have total/complete identity across all of their contexts. According to Murphy (2003), absolute synonyms must have the same meaning and be synonymous in all contexts. Palmer (1981) asserts that there are no absolute synonyms.

2) Near-synonyms are defined by DiMarco, Hirsts, and Stede (1993), as "words that are close in meaning, almost synonyms, but not quite; very similar, but not identical, in meaning; not fully intersubstitutable, but instead varying in their shades of denotations, connotations, implications, emphasis, or register." Murphy (2003), refers to the overlap of meaning and sense.

2.3 Criteria for Distinguishing Synonyms

In order to analyze the similarities and differences between the two synonyms "accurate" and "correct" in terms of meaning, degree of formality, collocation and

semantic preferences, this study used the data from the Oxford Advanced Learner's Dictionaries and the Corpus of Contemporary American English (COCA).

2.3.1 Senses of Meaning

Meaning is defined as “the thing or idea that a sound, word, sign, etc. represents.” Denotation and connotation are two distinct classifications for word meanings. Denotation refers to a word's direct or dictionary definition, reflecting the intended message of the speaker. In contrast, connotation deviates from the word's primary definition and may carry hidden or implied meanings.

2.3.2 Collections

Collocation is defined by the Oxford Advanced learner's Dictionaries as “a combination of words in a language that occurs very frequently and more frequently than would happen by chance.” According to the Cambridge Dictionary, collocation is the way words go together or are used in fixed connections. Collocation can be classified into two categories. Strong collocations appear when the relationship between the two words is tightly restricted or fixed. Weak collocations occur when a word may coexist with a large number of other words. For example, the noun “wish”, it can be used as an example of strong collocations due to the limited number of words that can collocate with it. In other hand, “big” was considered as a weak collocation. Various words can collocate with it.

2.3.3 Semantic preference

Semantic preference refers to semantic sets of collocates. Semantic preference is described by Stubbs (2002) as “a lexical set of frequently occurring collocates that share semantics features.” Cheng (2012) states that semantic preference is the tendency of lexical items to be restricted to particular semantic fields. According to Flowerdew (2012), semantic preference can help to organize words into themes or categories based on their meaning and semantic features.

2.3.4 Formality

Jackson & Amvela, (2000) states that the level of formality of the context is another stylistic variation that can be utilized to distinguish between near-synonyms. Formality pertains to the manner in which language is employed to depict contexts, circumstances, and the interplay between individuals engaged in communication. Learners can analyze the prevalence of certain word choices across various genres to gauge the level of formality. Certain sets of synonyms are more commonly employed in general or informal contexts, such as spoken language, TV/movies, or fiction, while others are better suited for formal usage, typically found in written texts such as websites, blogs, magazines, academic journals, and newspapers. Thongpan's study in 2022 examined the degree of formality by exploring the distribution of three synonyms: "far," "distant," and "remote." The study revealed that "far" is considered less formal due to its frequent usage in blogs, whereas "distant" and "remote" are more prevalent in magazines and academic contexts. Notably, "remote" is employed more frequently than "far." When comparing the degrees of formality among these three synonyms, it becomes apparent that "remote" is the most formal, followed by "distant," and "far" ranks as the least formal option, respectively.

2.4 Previous Related Studies

Several research studies have used corpus-based data from the Corpus of Contemporary American English (COCA) to evaluate the similarities and differences between synonyms based on meaning, collocations, formality, and grammatical patterns.

Rittisart (2021) conducted a study to differentiate between three synonymous verbs: obtain, gain, and receive. The researcher employed COCA and the Longman Dictionary of Contemporary English (LDOCE) as tools to analyze their grammatical patterns, degree of formality, and collocation. Despite their synonymous nature, these verbs exhibited some distinct grammatical patterns. In terms of formality, they predominantly appeared in written contexts with a high level of formality. Among the three, receive was found to be the most formal, followed by gain and obtain. Regarding collocation, only one noun collocation was shared among the three target verbs, and

they had a limited number of noun collocations in common. Thus, obtain, gain, and receive can be considered near synonyms that cannot be interchanged in all contexts.

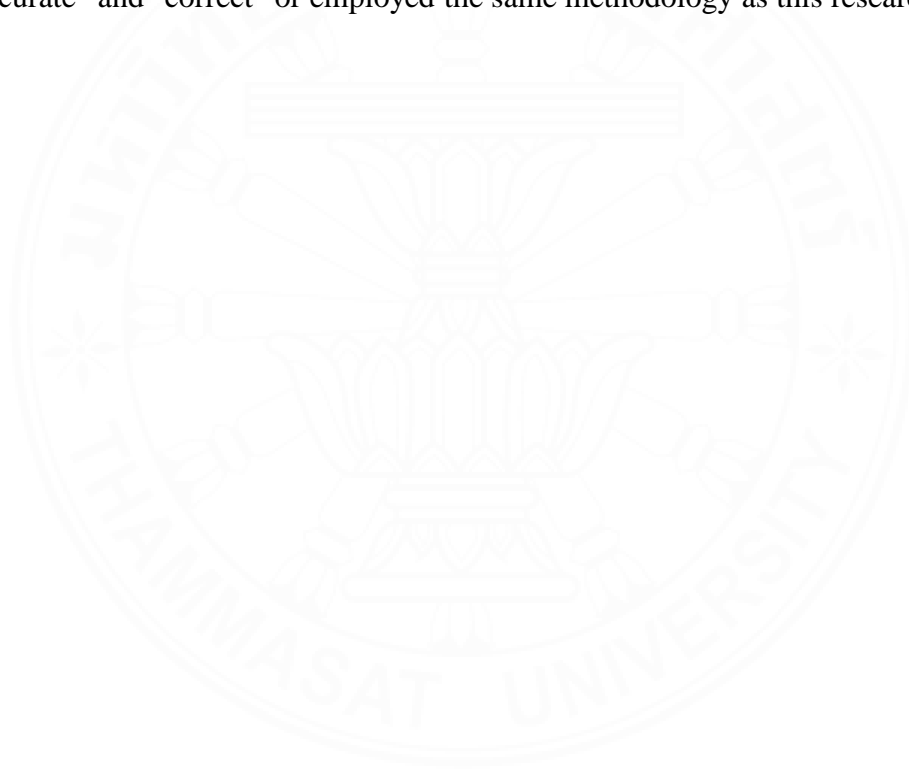
Waewsampun (2021) investigated the meanings, degree of formality, and collocations of three synonymous adjectives: real, genuine, and authentic. The researcher employed the Oxford Advanced Learner's Dictionary (OALD), the Longman Dictionary of Contemporary English (LDOCE), and COCA as tools. The results revealed that while the three synonyms shared a core meaning, they differed in their senses of meaning. In terms of formality, genuine and authentic were typically used in formal contexts with a high level of formality, whereas "real" had the lowest degree of formality. Concerning collocations, none of the three target words shared noun collocations with each other or with "real" and "authentic." However, "real" and "genuine" shared collocations such as sense, concern, interest, threat, and person, while "genuine" and "authentic" shared collocations such as dialogue and emotion. Due to these differences in meaning, formality, and collocations, these synonyms cannot be used interchangeably in all contexts.

Faemthaisong (2021) examined three synonymous adjectives: general, common, and typical, focusing on their meanings, noun collocations, grammatical patterns, and formality. Although "general" and "common" had closely related meanings, "common" was more commonly used in specific situations. While the three target words shared an identical meaning, their senses differed. Regarding noun collocations, some collocations were shared between each pair of synonyms, but none were common to all three. In terms of formality, "general" and "common" were used in both formal and informal contexts, while "common" was typically used in formal contexts. As for grammatical patterns, "common" exhibited a wider range of patterns compared to "general." Therefore, "general," "common," and "typical" are not absolute synonyms and cannot always be used interchangeably in certain contexts.

Sawat (2020) conducted a study to investigate the meanings, collocations, grammatical structures, and levels of formality between two synonymous words, say and mention, using dictionaries and COCA as sources. According to the findings from OALD and LDOCE, both words were defined as expressing something in words. However, "say" did not convey the meaning of a brief or short expression, whereas "mention" could. In terms of collocations, both "speak" and "mention" were frequently

used with nouns. Regarding grammatical patterns, although they shared core definitions, "say" was often used with clauses, while "mention" primarily occurred with noun phrases. In terms of formality, "mention" was more commonly used in formal situations compared to "say," although "say" frequently appeared in informal contexts. The study concluded that these two words were not absolute synonyms and could not always be used as substitutes for each other.

Although synonyms may share a core or similar meaning, there are often differences in aspects, such as formality, collocation, and semantic preference. To the best of the researcher's knowledge, none of the studies have researched the synonyms "accurate" and "correct" or employed the same methodology as this research.



CHAPTER 3

RESEARCH METHODOLOGY

This study's main goal was to compare and contrast the English words "accurate" and "correct." This chapter outlines the methodology and data gathering techniques used to answer the three study research questions. There are three key sections within it: (1) the target synonymous words, (2) research instrument, and (3) data collection and data analysis.

3.1 The Target Synonymous Words

The research target synonyms "accurate" and "correct" were selected because they appear frequently and are commonly used in English communication. The Oxford 3000, a list of the 3000 most essential English words, includes these two words. A team of linguists and experienced teachers selected the Oxford 3000 keywords as the terms that should be given priority in vocabulary study due to their significance and utility (The Oxford 3000, n.d.). According to the Oxford Advanced Learner's Dictionary (online version), since two target words have the same core meaning, it is important to understand and analyze their variations in usage. The Corpus of Contemporary American English (COCA) was used in this study to investigate the similarities and differences between the two target synonyms. The COCA contains much data about the use of English in authentic contexts across eight genres: spoken language, fiction, TV/movies, websites, blogs, magazines, newspapers, and academic journals, which are very useful to explore the frequency and distribution across different genres and to investigate the collocation and semantic preference of these two words.

3.2 Research Instruments

3.2.1 The Oxford Advanced Learner's Dictionaries (Online Version)

The Oxford Advanced Learner's Dictionaries (online version) is a well-known and easily accessible dictionary. It provides definitions that learners can understand, example sentences demonstrating language usage, pronunciations with both American and British accents and some grammatical patterns, which is very useful for learners.

3.2.2 Corpus of Contemporary American English (COCA)

More than a billion words are included in the Corpus of Contemporary American English (COCA), including spoken languages, TV and movies, academic journals, newspapers, magazines, blogs and websites. It is a sizable dataset that offers English language in real-world settings and covers a variety of genres. It includes numerous examples based on authentic English. The COCA function is very user-friendly. The collocation feature, one of several useful features, allows the user to explore the data and evaluate various collocations of the target term based on frequency and MI score.

3.3 Data Collection and Data Analysis

This section provides three main processes for distinguishing the similarities and differences between the two target synonyms in terms of meaning, degree of formality, collocations and semantic preference. Both qualitative and quantitative approaches of analysis were used in this study. The frequency of each term and the co-occurrence of collocations with each word are compared using a quantitative approach. In conducting qualitative research, the relationship between words and a category of terms that have similar semantics is considered.

3.3.1 Data Collection

To examine the similarities and differences in meaning between synonyms, “accurate” and “correct”, the Oxford Advanced Learner's Dictionary was used by searching on the website: <https://www.oxfordlearnersdictionaries.com/>.

The Corpus of Contemporary American English (COCA), <https://www.english-corpora.org/coca/> was used to determine formality and collocations. To investigate the degree of formality as determined by the frequency and distribution across genres, the researcher entered the target words in the search box and clicked on the chart function; the results in terms of frequency in each genre appeared. In order to investigate the noun and adverb collocation, the synonyms were put into the search box to obtain the word functions and collocations. This study collected the top 30 noun and adverb collocations that occurred the most with the target synonyms, while noun and adverb collocations with mutual information or MI scores of less than 3 were removed. The MI score is a statistical test for collocational strength. Hunston (2002) states that only MI scores of 3

or higher should be considered to be of importance for any given collocation. The noun and adverb collocations which were collected were assigned to themes based on their similarity in meaning.

3.3.2 Data Analysis

The following procedures were utilized to look into the similarities and differences in meaning, formality, collocations and semantic preferences between the synonyms “accurate” and “correct” in order to provide answers to the three study questions.

First, the definitions and examples of word usage from the Oxford Advanced Learner's Dictionary were investigated to determine the similarities and differences in meaning. Second, COCA data on the distribution of two synonym words across eight genres were collected in order to examine the degree of formality between the two synonyms. The COCA results of distribution across eight genres are displayed in a chart with the frequency of each target word in each section, as well as the frequency of those words per million. As for collocations, the top 30 noun and adverb collocations were represented based on frequency. Finally, after the collocations were extracted, the nouns and adverbs that frequently co-occur with the synonyms were classified into the same themes based on semantic preferences or similar meaning to determine the similarities and differences between “accurate” and “correct.”

CHAPTER 4

RESULTS AND DISCUSSION

In order to deal with the three research questions given below, this chapter compares and contrasts the two synonyms accurate and correct using information from the Corpus of Contemporary American English (COCA) and Oxford Advanced Learner's Dictionary (OALD).

RQ. 1: What are the similarities and differences between the synonyms accurate and correct in terms of meaning?

RQ. 2: What are the similarities and differences between the synonyms accurate and correct in terms of the degree of formality determined by the frequency and distribution across genres?

RQ. 3: What are the similarities and differences between the synonyms accurate and correct in terms of collocations and semantic preference?

4.1 Similarities and Differences in Meaning

Table 4.1

Definition of Accurate and Correct from the Oxford Advanced Learner's Dictionary

Synonym	Definition	Example
Accurate	1. correct and true in every detail	- The film is not historically accurate
	2. able to give completely correct information or to do something in an exact way	- My watch is not very accurate.
Correct	1. accurate or true, without any mistakes	- Please check that these details are correct. - Do you have the correct time?
	2. right and suitable, so that something is done as it should be done	- I think you've made the correct decision.

The words “accurate” and “correct” share a core meaning of ‘true,’ as seen in table 4.1, which contains the definition and examples of an adjective for two synonyms from the Oxford Advanced Learner's Dictionary. However, they have some differences in the senses in which they are used; “accurate” can refer to something that is precise or exact, whereas “correct” refers to something that is right.

The findings showed that although these two synonyms share the same core meaning, they differ in other sense, making it impossible for them to be used interchangeably in all situations.

4.2 Degree of Formality

The Corpus of Contemporary American English (COCA) was used as the instrument to explore the formality as determined by the frequency and distribution across eight different genres: spoken language, newspapers, magazines, academic writings, fiction, TV and movie subtitles, blogs, and other websites. The results are presented in table 4.2.

Table 4.2

Overall Frequency of Accurate and Correct Across the Eight Genres from COCA

Genre	<i>Accurate</i>		<i>Correct</i>	
	Frequency	Per Million	Frequency	Per Million
Academic	6,747	56.32	9,620	80.31
Blogs	5,822	45.27	12,946	100.66
Fiction	1,163	9.83	3,877	32.77
Magazines	4,423	35.08	5,339	42.34
News	2,879	23.65	3,706	30.44
Spoken language	3,359	26.63	14,886	118.02
TV/Movies	1,182	9.23	6,167	48.15
Websites	5,529	44.5	11,718	94.31
Total	31,104		68,259	

Table 4.2 shows that “correct” occurs with the highest frequency (68,259 tokens), whereas accurate occurs with a frequency of 31,104 tokens, which is half that of “correct”.

Table 4.3

Distribution of Accurate and Correct across Eight Genres according to Frequency

Genre	<i>Accurate</i>		<i>Correct</i>		
	Frequency	Per Million	Genre	Frequency	Per Million
Academic	6,747.00	56.32	Spoken	14,886.00	118.02
Blogs	5,822.00	45.27	Blogs	12,946.00	100.66
Websites	5,529.00	44.5	Websites	11,718.00	94.31
Magazines	4,423.00	35.08	Academic	9,620.00	80.31
Spoken	3,359.00	26.63	TV/Movie	6,167.00	48.15
NEWS	2,879.00	23.65	Magazines	5,339.00	42.34
Fiction	1,163.00	9.83	Fiction	3,877.00	32.77
TV/Movie	1,182.00	9.23	NEWS	3,706.00	30.44
Total	31,104.00		Total	68,259.00	

In Table 4.3, the data reveals distinct usage patterns for the synonyms "accurate" and "correct." The frequency analysis demonstrates that "accurate" is predominantly associated with written English, with the highest occurrences in academic texts, blogs, and websites. Specifically, it appeared 6,747 times (56.32 per million) in academic contexts, 5,822 times (45.27 per million) in blogs, and 5,529 times (44.50 per million) in websites. On the other hand, "correct" was most frequently utilized in spoken English, with a frequency of 14,886 tokens (118.02 per million), followed by blogs (12,946 tokens) and websites (11,718 tokens). Notably, "correct" occurred more frequently than "accurate" in every genre, approximately twice as often. These findings suggest that "correct" is more commonly employed in general English and academic writing compared to "accurate."

When considering the degree of formality, "accurate" was found to be most prevalent in academic texts, which typically adhere to written language norms and formal contexts. Conversely, "correct" appeared more frequently in spoken English, indicating its typical usage in informal contexts.

4.3 Collocations

In this section, the two synonyms "accurate" and "correct" are adjectives; therefore, they commonly occurred with nouns and adverbs. Even though they share a similar core meaning, they may not share the same collocations. This study gathered the top 30 frequency noun and adverb collocations that occur with these two synonyms by selecting those with MI scores equal or higher to 3, which is the significant value for a collocational relationship in order to investigate the similarity and difference in terms of collocation between the synonym "accurate" and "correct".

4.3.1 Noun Collocations

The top 30 noun collocations with accurate and correct are listed in Table 4.4.

Table 4.4

Top-30 Frequency List of Noun Collocation of Accurate and Correct from COCA

Rank	<i>Accurate</i>			<i>Correct</i>		
	Noun Collocates	Frequency	MI Score	Noun Collocate	Frequency	MI Score
1	information	1,546	4.3	answer	2,018	5.11
2	description	551	5.71	response	812	3.74
3	reporting	538	6.76	interpretation	270	4.18
4	data	499	3.09	percentage	222	3.39
5	picture	486	3.55	spelling	187	6.02
6	assessment	420	4.91	assumption	168	3.54
7	prediction	418	6.31	diagnosis	157	4.13
8	statement	333	3.44	pronunciation	147	7.12
9	representation	327	5.68	prediction	141	3.9

Rank	<i>Accurate</i>			<i>Correct</i>		
	Noun Collocates	Frequency	MI Score	Noun Collocate	Frequency	MI Score
10	measurement	318	5.74	sequence	131	3.32
11	estimate	299	5.26	calculation	116	4.19
12	diagnosis	266	5.73	usage	103	4.26
13	count	255	4.95	hypothesis	101	3.44
14	reading	254	3.96	grammar	89	4.86
15	account	220	3.11	proportion	83	3.07
16	measure	205	3.12	suspicion	74	3.45
17	portrayal	185	7.27	instinct	63	3.24
18	map	177	3.6	punctuation	62	6.3
19	reflection	175	4.86	terminology	58	4.93
20	poll	154	3.32	premise	56	3.06
21	depiction	131	6.65	guess	55	3.06
22	forecast	120	5.54	posture	53	3.91
23	perception	119	3.54	classification	50	3.42
24	observation	115	3.54	digit	44	4.03
25	identification	101	4.28	assertion	43	3.04
26	clock	87	3.58	password	38	3.03
27	translation	81	4.24	dosage	38	5.04
28	interpretation	79	3.24	hunch	32	4.62
29	statistics	76	3.38	verb	25	3.19
30	predictor	75	5.12	intuition	25	3.42

Based on their frequency and MI scores (≥ 3), Table 4.4 lists the top 30 noun collocations that frequently occur with the two target words. As can be seen there are three noun collocates that are shared by the two target synonyms, i.e., ‘prediction’, ‘diagnosis’ and ‘interpretation’.

To evaluate the similarities and differences of noun collocates, the top 30 noun collocations of each word were examined by grouping them into the same semantic theme based on semantic features and meaning. The results are displayed in Table 4.5 and 4.6.

Table 4.5

Semantic Preference of Noun Collocates of Accurate from COCA

Semantic Categories	Noun Collocates of Accurate
1. information	information, description, data, reporting, account, statement
2. linguistic	translation, interpretation, reading
3. medical	diagnosis
4. number	assessment, measurement, statistics, poll, count
5. object	picture, map, clock
6. possibility	prediction, forecast, estimate
7. sense of sight	picture, perception, representation, reflection, observation, depiction
8. miscellaneous	identification

Table 4.5 shows that the noun collocations of accurate were categorized into eight themes; information, linguistic, medical, number, object, possibility, sense of sight and miscellaneous. “Accurate” is mostly related to the theme of information and sense of sight, which contains six noun collocations. Information, description, data, reporting, account, and statement are all included in the theme of information. Regarding the theme of the sense of sight, this includes picture, perception, representation, reflection, observation, depiction. The following are some examples of noun collocations derived from COCA concordance lines.

1. A well-maintained article will have all kinds of links to follow to gather **accurate information**.
2. Provide sufficient context to allow for **accurate interpretation** of the data.
3. Flexible endoscopy allows the **accurate diagnosis** of TIE in more than 90% of cases.

4. You have to have an **accurate count** of how much property was damaged.
5. If you have a detailed and **accurate map**, finding what you want is going to be an easy and efficient.
6. In fact, the most **accurate predictor** of a student's achievement in school is not income or social status.
7. Biased data is not an **accurate reflection** of reality.
8. **Accurate identification** of Plasmodium infections is critical for administration of a targeted therapy.

Table 4.6

Semantic Preference of Noun Collocates of Correct from COCA

Semantic Categories	Noun Collocates of <i>Correct</i>
1. linguistic	pronunciation, usage, grammar, punctuation, terminology, verb, spelling, interpretation
2. medical	diagnosis, dosage
3. number/amount	percentage, calculation, digit, password, proportion
4. research	assumption, prediction, sequence, hypothesis, premise
5. response	answer, response
6. sense	instinct, intuition, hunch, guess, suspicion
7. miscellaneous	posture, assertion, classification

As shown in Table 4.6, noun collocations of “correct” were categorized into seven categories, including linguistic, medical, number/amount, research, respond, sense and miscellaneous. The majority of nouns that co-occur with “correct” are associated linguistic theme, including words, e.g., pronunciation, usage, grammar, punctuation, terminology, verb, spelling and interpretation. Concordance lines demonstrate that these nouns collocate with “correct.” They are shown in the examples below.

1. The latter is undoubtedly the **correct interpretation** here.
2. They will help get you a **correct diagnosis** because there are many kinds of migraines.

3. The **correct percentage** would be 38 percent
4. Accepting a hypothesis does not guarantee that it is the **correct hypothesis!**
5. I mean, if there's no **correct answer**, it must be a trick question.
6. You had a **correct instinct** that the statement was overbroad but no particular knowledge or documentation to bring forward.
7. They work together with the back muscles to maintain **correct posture** and stabilize the spine.

As shown in Tables 4.5 and 4.6, the two synonyms, accurate and correct, differ in a variety of semantic domains, including information, object, possibility, sense of sight, research, response, and sense. However, they have overlapping semantic categories, such as linguistic, medical, and numbers.

4.3.2 Adverb Collocations

The adverb collocations are presented in the table 4.7.

Table 4.7

Top-30 Frequency List of Adverb Collocations of Accurate and Correct from COCA

Rank	<i>Accurate</i>			<i>Correct</i>		
	Adv. collocate	Freq.	MI Score	Adv. collocate	Freq	MI Score
1	more	4,723	3.92	politically	2,368	3.92
2	less	429	3.17	absolutely	936	3.17
3	historically	279	6.15	technically	215	6.15
4	highly	243	3.82	grammatically	180	3.82
5	entirely	228	4.21	anatomically	120	4.21
6	fairly	226	4.63	partially	111	4.63
7	extremely	133	3.36	morally	110	3.36
8	reasonably	131	5.51	factually	104	5.51
9	scientifically	80	6.41	historically	59	6.41
10	perfectly	77	3.02	environmentally	57	3.02
11	technically	77	4.58	scientifically	29	4.96

Rank	<i>Accurate</i>			<i>Correct</i>		
	Adv. collocate	Freq.	MI Score	Adv. collocate	Freq	MI Score
12	factually	71	8.09	logically	24	3.37
13	remarkably	69	4.96	ergonomically	24	3.37
14	incredibly	60	3.37	ethically	22	4.12
15	surprisingly	46	3.37	objectively	20	4.81
16	sufficiently	37	4.12	ecologically	18	5.23
17	amazingly	34	4.81	mathematically	16	3.05
18	medically	26	5.23	theologically	16	3.93
19	strictly	25	3.05	semantically	16	7.9
20	uncannily	24	7.9	syntactically	16	3.58
21	wholly	20	3.58	ideologically	14	5.55
22	eerily	20	5.55	biologically	10	3.69
23	remotely	19	3.69	demographically	10	3.01
24	culturally	16	3.01	constitutionally	9	4.21
25	empirically	13	4.39	educationally	9	4.39
26	herein	13	4.43	procedurally	9	4.43
27	painfully	10	3.07	directionally	9	3.07
28	clinically	10	3.9	intrinsically	8	3.9
29	geographically	10	4.05	biblically	8	4.05
30	plenty	10	4.51	philosophically	7	4.51

The top 30 adverb collocations between the two target words are shown in Table 7 regarding their frequency and MI scores (≥ 3). As can be seen, the adverb collocations "historically", "scientifically" "technically" and "factually" share the two target synonyms.

In order to determine the similarities and differences of the adverb collocates, the top 30 adverb collocates of each word were studied by grouping them into the same semantic theme based on semantic features and meaning. Tables 8 and 9 present the results.

Table 4.8*Semantic Preference of Adverb Collocates of Accurate from COCA*

Semantic Categories	Adverb Collocates of Accurate
1. academic	culturally, historically, geographically
2. fact	clinically, empirically, factually, medically, scientifically, technically
3. feeling	surprisingly, amazingly, incredibly, remarkably, painfully, eerily, uncannily
4. level/degree	highly, extremely, plenty, fairly, reasonably, wholly, entirely, remotely, more, less
5. miscellaneous	perfectly, sufficiently, strictly, herein

According to Table 4.8, the adverb collocations of “accurate” were separated into five categories: academic, fact, feeling, level/degree and miscellaneous. “Accurate” has the highest number of collocations in the theme of level/degree, including words, e.g., highly, extremely, plenty, fairly, reasonably, wholly, entirely, remotely, more, less. The examples of adverb collocations based on COCA concordance lines are shown below.

1. It's more schematic than **geographically accurate**.
2. University officials can make some **surprisingly accurate** predictions about who will flourish and who will flounder.
3. This could have provided **more accurate** data as to how these processes work as the comet interacts with the Sun.
4. Presentation of material by a physician is **scientifically accurate**.
5. Flu tests are not **perfectly accurate**.

Table 4.9*Semantic Preference of Adverb Collocates of Correct from COCA*

Semantic Categories	Adverb Collocates of <i>Correct</i>
1. academic	anatomically, biologically, educationally, environmentally, ergonomically, historically, mathematically, philosophically, theologically, grammatically, semantically, syntactically
2. fact	factually, objectively, scientifically, technically
3. level/degree	absolutely, partially
4. moral	ethically, morally
5. political	constitutionally, politically
6. thought	ideologically, logically
7. miscellaneous	biblically, demographically, procedurally, directionally, intrinsically

Table 4.9 displays the adverb collocations of “correct” that were divided into seven groups: academic, fact, level/degree, moral, political, thought and miscellaneous. The category with the most adverb collocations is academic, which includes anatomically, biologically, educationally, environmentally, ergonomically, historically, mathematically, philosophically, theologically, grammatically, semantically, syntactically. The following are examples of adverb collocations based on COCA concordance lines.

1. Sounds weird to you because it is not **grammatically correct**.
2. I couldn't tell you if it is truly **scientifically correct**, but for the purpose of the book it makes sense.
3. Highly idealistic individuals believe that **ethically correct** actions will consistently produce desirable outcomes.
4. It may have been the most **politically correct** terrorism bust in history.
5. he gave it a low grade because it was not **ideologically correct**.
6. This article is **absolutely correct** that the poor make less money and pay more taxes.

7. In contrast, takes the view that the transition from authoritarianism to **procedurally correct** elections produces only hybrid regimes.

The two synonyms, “accurate” and “correct,” differ in some semantic categories such as feeling, moral, political, and thought, but they share common themes in others such as academic, fact, and level/degree as shown in Tables 4.8 and 4.9.

According to the study, the terms “accurate” and “correct” are near synonym since they share core meanings, some noun and adverb collocations, and some semantic categories. However, they have some differences in sense of meaning and collocation; therefore, they cannot be used interchangeably in every context.



CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

This chapter covers (5.1) a summary of the study, (5.2) a summary of the findings, (5.3) the conclusion, (5.4) the implications and (5.5) limitations and recommendations for future research

5.1 Summary of the study

This section aims to explore the similarities and differences between the synonyms “accurate” and “correct” based on several criteria, including meaning, formality level, collocations, and semantic preferences.

5.1.1 Objective of the Study

The objective of this study was to investigate the similarities and differences between “accurate” and “correct,” with a focus on their meaning, formality, collocations and semantic preferences, using the Oxford Advanced Learner's Dictionary (online version) and the Corpus of Contemporary American English (COCA) as the main research tools.

5.1.2 Selected Synonyms

The two synonymous adjectives, “accurate” and “correct,” are included in the Oxford 3000 list, which consists of high-frequency words commonly used in both spoken and written English. The Oxford 3000 keywords were carefully selected by language experts and experienced teachers as essential vocabulary items for prioritized study (The Oxford 3000, n.d.). The referential meanings of these words were obtained from the online version of Oxford Learner's Dictionaries. To analyze the formality level and collocations, data was collected from the Corpus of Contemporary American English (COCA). The researcher utilized COCA to investigate the frequency and distribution of each synonym across eight genres, including spoken language, fiction, TV/movies, websites, blogs, magazines, newspapers, and academic journals. Collocates between the two synonyms were identified using COCA, focusing on the

top 30 noun and adverb collocations with MI (Mutual Information) scores of 3 or above. These collocations were then classified according to semantic preference to compare and contrast the usage of the two synonyms. Concrete examples of concordance lines extracted from COCA were collected to provide clarity on how the nouns and adverbs associated with the synonyms in collocation.

5.2 Summary of the findings

The findings in three areas, similarities and differences in meaning, the degree of formality, collocations and semantic preferences based on data from the COCA and the Oxford Advance Learner's Dictionaries (online version) are summarized, as follows.

5.2.1 Definition of Accurate and Correct from the Oxford Advanced Learner's Dictionary

According to the online Oxford Advanced Learner's Dictionary, the target words accurate and correct share a core meaning of 'true', with "accurate" defined as correct and true in every detail and "correct" defined as accurate or true, without any mistakes. However, the contexts in which they are used differ; "accurate" might relate to something that is exact, whereas "correct" refers to something that is right. As a result, they are unable to be used interchangeably in all contexts.

5.2.2 Degree of formality

According to statistics from the Corpus of Contemporary American English (COCA), the word "correct" appears the most frequently (68,259 tokens), whereas "accurate" occurs half as frequently (31,104 tokens). The frequency and distribution across eight different genres showed that "accurate" is more associated with written English, as it occurs most frequently in academic texts, blogs and websites with a frequency of 6,747 tokens (56.32 per million), 5,822 tokens (45.27 per million) and 5,529 tokens (44.50 per million), respectively. Since formal English is intended to be used in academic writing, it can be considered that "accurate" is related to formal contexts and has a high degree of formality. Contrarily, the word "correct" is related to

informal contexts, as it occurs most frequently in spoken English with a frequency of 14,886 tokens (118.02 per million).

5.2.3 Collocations

Analysis of the COCA corpus data provided valuable insights into the collocations of the synonyms “accurate” and “correct.” The top 30 noun and adverb collocations, selected based on a MI score of 3 or higher, shed light on their usage patterns.

Regarding noun collocations, three common collocations were identified: "prediction," "diagnostic," and "interpretation," which were associated with both synonyms. Additionally, the top 30 frequency nouns were categorized based on semantic preference. For “accurate,” the noun collocations fell into eight categories, including information, linguistic, medical, number, object, possibility, sense of sight, and miscellaneous. On the other hand, the noun collocations for “correct” were classified into seven categories, including linguistic, medical, number/amount, research, response, sense, and miscellaneous. Notably, there were overlapping semantic categories such as linguistic, medical, and number.

Moving on to adverb collocations, "historically," "scientifically," "technically," and "factually" were shared by both synonyms. The adverb collocations for “accurate” were grouped into five categories: academic, fact, feeling, level/degree, and miscellaneous. Similarly, the adverb collocations for "correct" were divided into seven categories: academic, fact, level/degree, moral, political, thought, and miscellaneous. Common themes emerged in both synonyms, including academic, fact, and level/degree.

From the analysis, it is evident that there are several distinct noun and adverb collocations between these two synonyms. However, they also share certain collocations, suggesting that although they are not absolute replacements, in some contexts, they can be used interchangeably.

5.3 Conclusion

The two synonyms are regarded as near-synonyms, according to the study’s findings. They are distinguished by the criteria examined in this study. In terms of

degree of formality, COCA data reveals that “accurate” occurs with a high level of formality because it tends to be used in written contexts, whereas “correct” is the most frequently used in spoken language, which is considered as informal. However, when comparing two synonyms, “correct” has a higher frequency in written contexts than “accurate”, indicating that “correct” is more commonly used in formal and informal contexts. With regard to collocations, focusing on the noun and adverb collocations, these two synonyms share three common noun collocations: ‘prediction’, ‘diagnosis,’ and ‘interpretation’ and four adverb collocations ‘historically’, ‘scientifically’, ‘technically’ and ‘factually’. Even though two synonyms have the same basic meaning of true, their senses of meaning and some collocations are different; therefore, they cannot be substituted in all situations because they each have their own distinct features.

5.4 Implications

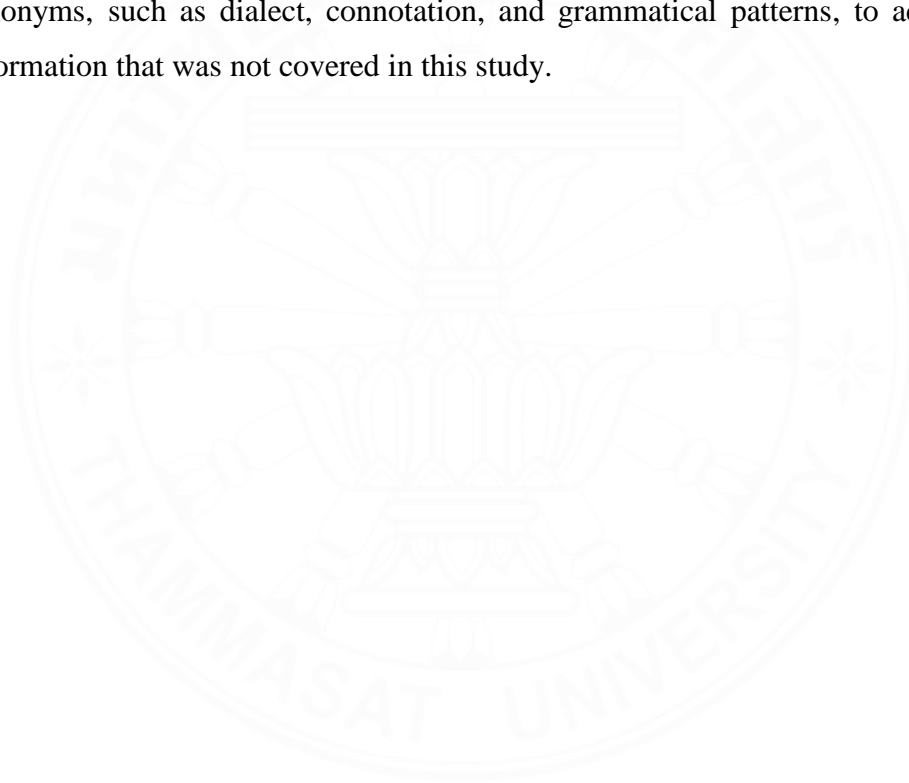
Non-native English speakers, particularly those learning English as a second language or as a foreign language, often face challenges in accurately utilizing near synonyms. The findings of this research highlight the limitations of relying solely on dictionaries to determine word meanings and choose the most appropriate synonym for a given context. Dictionaries often fail to provide learners with sufficient information about the similarities and differences between near synonyms. It is important to recognize that synonyms can vary in certain aspects, even when they appear to have similar meanings or share some collocations. To address this issue, learners can benefit from utilizing the COCA corpus as a resource. The COCA corpus offers valuable insights into the frequency of word usage across various genres and helps learners identify whether certain synonyms are more formal, informal, or somewhere in between. By examining the differences and similarities between each synonym using this information, learners can make more informed choices in selecting the most suitable word for a specific situation. Furthermore, utilizing COCA allows learners to observe how the target vocabulary is employed in authentic contexts through concordance lines. This exposure to real-world usage and everyday English enhances the effectiveness of learning compared to solely relying on dictionary examples.

In conclusion, non-native English speakers can enhance their understanding and usage of near synonyms by utilizing the COCA corpus. This approach provides

valuable information on frequency, formality, and usage contexts, empowering learners to make more accurate and appropriate word choices in their language production.

5.5 Limitations and Recommendations for Future Research

This study's focus was on the Corpus of Contemporary American English (COCA). Consideration should also be given to other well-known native-speaker corpora, such as the British Nation Corpus (BNC), which is a collection of words that includes authentic examples of British English that have been written and spoken in a variety of contexts. Future research could also employ additional criteria to identify synonyms, such as dialect, connotation, and grammatical patterns, to acquire more information that was not covered in this study.



REFERENCES

- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and us.*, Cambridge University Press.
- Cheng, W. (2012). *Exploring corpus linguistics*. Language in action. Routledge.
- Cruse, D. A. (2000). *Meaning in language: An introduction to semantics and pragmatic*. OUP Oxford.
- DiMarco, C., Hirst, G., & Stede, M. (1993). *The semantic and stylistic differentiation of synonyms and near-synonyms*. AAAI Spring Symposium on Building Lexicons for Machine Translation, Stanford, CA.
- Faemthaisong, P. (2021). *A corpus-based study of English synonyms: General, common, and typical* [Master's thesis, Thammasat University].
- Flowerdew, L. (2012). *Corpora and language education*. Palgrave Macmillan.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
- Jackson, H., & Amvela, E. (2000). *Words, meaning and vocabulary: Meaning relations*. Athenaum Press.
- Longman Dictionary of Contemporary English: LDOCE. (n.d.). Available online at <https://www.ldoceonline.com>.
- Murphy, M. L. (2003). *Semantic relation 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.
- Oxford Advance Learner's Dictionary. (n.d.). Available online at <https://www.oxfordlearnersdictionaries.com>.
- Palmer, F. R. (1997). *Semantics*. Cambridge University Press.
- Palmer, F. R. (1981). *Semantics*. (2nd ed.). Cambridge University Press.
- Rittisart, R (2021). *A corpus-based analysis of English synonym verbs: Obtain, gain, and receive* [Master's thesis, Thammasat University].
- Rozakis, E. L. (2011). *Vocabulary for dummies*. Wiley.
- Sawat, T. (2020). *A corpus-based analysis of English synonymous verbs: Say and mention* [Master's thesis, Thammasat University].

- Stubbs, M. (2002). Two quantitative methods of studying phraseology in English. *International Journal of Corpus Linguistics*, 7, 215-244. <https://doi.org/10.1075/ijcl.7.2.04stu>.
- Taylor, J. R. (2002). Near synonyms as co-extensive categories: 'High' and 'tall' revisited. *Language Sciences*, 25, 263-284.
- Thongpan, N. (2021). *A Corpus-based study of English synonyms of the adjective 'Far' and 'Distant', and 'Remote'* [Master's thesis, Thammasat University].
- Thornbury, S. (2002). *How to teach vocabulary*. Longman.
- Waewsampun, J. (2021). *A corpus-based analysis of English synonyms: Real, genuine, and authentic* [Master's thesis, Thammasat University].

