



**A CORPUS-BASED STUDY OF THE NEAR-SYNONYMS:
BEAUTIFUL, ATTRACTIVE, AND GORGEOUS**

BY

KOBKARN JARANIPATT

**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 THE NEAR-SYNONYMS:
BEAUTIFUL, ATTRACTIVE, AND GORGEOUS**

BY

KOBKARN JARANIPATT



**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

KOBKARN JARANIPATT

ENTITLED

A CORPUS-BASED STUDY OF THE NEAR-SYNONYMS: BEAUTIFUL,
ATTRACTIVE, AND GORGEOUS

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

on July 30, 2024

Chairman

Supong Tangkiengsirisin

(Associate Professor Supong Tangkiengsirisin, Ph.D.)

Member and Advisor

Supakorn Phoocharoensil

(Associate Professor Supakorn Phoocharoensil, Ph.D.)

Director

Supakorn Phoocharoensil

(Associate Professor Supakorn Phoocharoensil, Ph.D.)

Independent Study Title	A CORPUS-BASED STUDY OF THE NEAR-SYNONYMS: BEAUTIFUL, ATTRACTIVE, AND GORGEOUS
Author	Kobkarn Jaranipatt
Degree	Master of Arts
Major Field/Faculty/University	English Language Teaching Language Institute Thammasat University
Independent Study Advisor	Associate Professor Supakorn Phoocharoensil, Ph.D.
Academic Year	2023

ABSTRACT

This corpus-based study aims to investigate three near-synonyms—*beautiful*, *attractive*, and *gorgeous*—regarding their distribution across genres and their noun and adverb collocations in relation to semantic preference. Data were drawn from two well-known online dictionaries and the Corpus of Contemporary American English (COCA). The findings reveal that *beautiful* and *gorgeous* predominantly occur in spoken genres (e.g., TV and movie, subtitles, fiction, and spoken), while *attractive* is more prevalent in written genres (e.g., magazines and blogs). Interestingly, the analysis demonstrated that while all three adjectives shared some noun collocates and semantic preferences related to humans and settings, *attractive* had broader usage, extending into business fields. The similarities and differences in meanings and semantic fields suggest that they are more accurately classified as near-synonyms, which are not entirely interchangeable. The comprehensive investigation provides valuable insights for EFL instruction, emphasizing the importance of utilizing multiple resources to discern subtle differences among near-synonyms and promote their effective use in communication.

Keywords: synonyms, near-synonyms, collocations, semantic preference, corpus

ACKNOWLEDGEMENTS

As I complete this independent study, I am filled with gratitude for those who have supported me throughout this journey.

First and foremost, I would like to express my deepest gratitude to my advisor, Associate Professor Dr. Supakorn Phoocharoensil, for providing invaluable guidance, support, and patience throughout the process of completing this independent study. His expertise and feedback have been instrumental in shaping both this study and my academic growth.

Secondly, I extend my sincere appreciation to all the instructors at the Language Institute of Thammasat University. The knowledge and experience in language teaching that you have imparted throughout my academic journey have been invaluable. Your dedication to education has inspired me to strive for excellence in my career path.

Thirdly, I would like to thank all my classmates, with special gratitude to Azin, Chayanid, and Kamomchanok, who have always been great teammates and supported each other from beginning to end of this program. The friendship and support from all of you have made this experience truly memorable and enjoyable.

Lastly, I extend my gratitude to my family and friends for their encouragement throughout this journey. Your patience, understanding, and belief in me have been a source of strength and motivation, especially during challenging times.

Kobkarn Jaranipatt

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	2
1.3 Research questions	2
1.4 Definitions of key terms	2
1.5 Scope of the study	3
1.6 Significance of the study	3
1.7 Organization of the study	4
CHAPTER 2 REVIEW OF LITERATURE	
2.1 The concept of corpora	5
2.1.1 The Corpus of Contemporary American English (COCA)	5
2.2 The concept of synonymy	7
2.3 Criteria for distinguishing synonyms	7
2.3.1 Senses of meaning	8
2.3.2 Degree of formality	9
2.3.3 Collocations	9
2.3.4 Semantic preference	10

2.4 Previous studies	10
----------------------	----

CHAPTER 3 RESEARCH METHODOLOGY

3.1 Target near-synonyms	14
3.2 Data collection	15
3.3 Data analysis	16

CHAPTER 4 RESULTS

4.1 Meanings	17
4.2 Frequency and distribution of the synonyms across genres	19
4.3 Collocations	21
4.3.1 Noun collocations	21
4.3.2 Adverb collocations	29

CHAPTER 5 DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary of findings	36
5.1.1 Meanings	36
5.1.2 Genres distribution of the target words	36
5.1.3 Noun collocations	37
5.1.4 Adverb collocations	37
5.2 Discussion	37
5.2.1 Meanings	38
5.2.2 Genres distribution of the target words	39
5.2.3 Noun collocations	39
5.2.4 Adverb collocations	40
5.3 Conclusion	42
5.4 Implications	42

	(5)
5.5 Limitations	43
5.6 Recommendations	44
REFERENCES	45

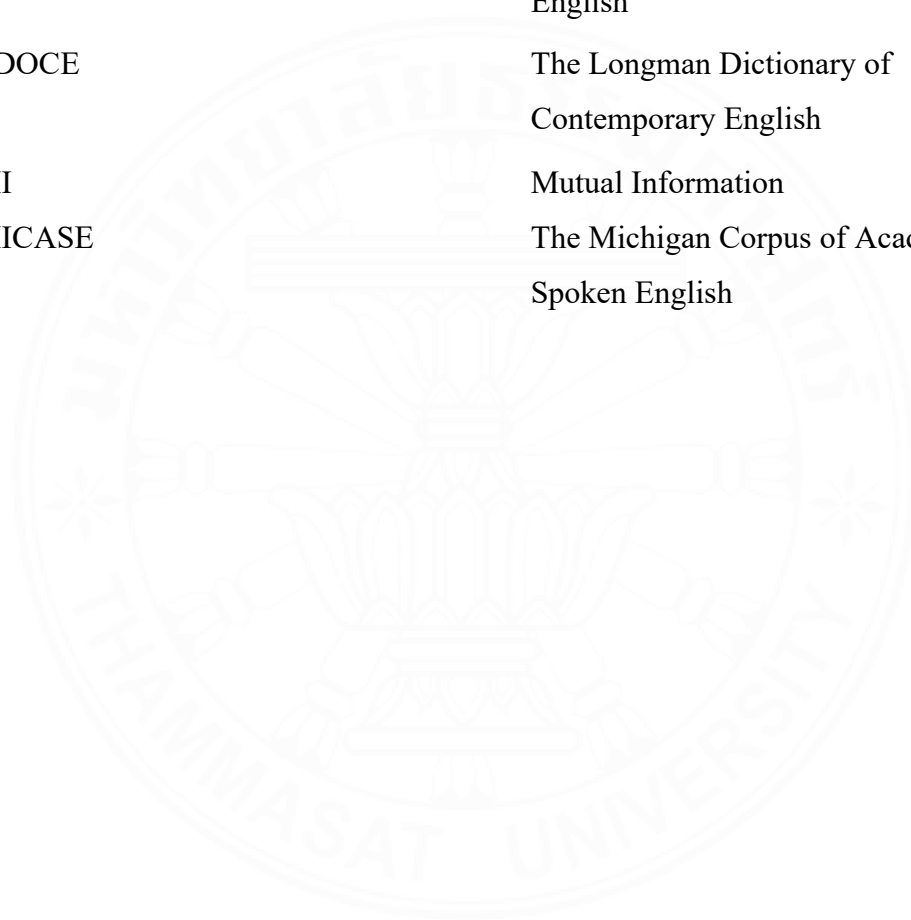


LIST OF TABLES

Tables		Page
1. Meanings of “beautiful,” “attractive,” and “gorgeous” from the Longman dictionary of contemporary English online (2024) and Merriam-Webster dictionary online (2024)		17
2. The overall frequency of “beautiful,” “attractive,” and “gorgeous” across eight genres from COCA		19
3. Distribution of “beautiful,” “attractive,” and “gorgeous” according to frequency		20
4. Top-30 frequency list of noun collocates of “beautiful,” “attractive,” and “gorgeous” in COCA		21
5. Semantic preference of noun collocates of “beautiful.”		24
6. Semantic preference of noun collocates of “attractive.”		26
7. Semantic preference of noun collocates of “gorgeous.”		27
8. Top-30 frequency list of adverb collocates of “beautiful,” “attractive,” and “gorgeous” in COCA		29
9. Semantic preference of adverb collocates of “beautiful.”		31
10. Semantic preference of adverb collocates of “attractive.”		32
11. Semantic preference of adverb collocates of “gorgeous”		34

LIST OF ABBREVIATIONS

Symbols/Abbreviations	Terms
BNC	The British National Corpus
COCA	The Corpus of Contemporary American English
LDOCE	The Longman Dictionary of Contemporary English
MI	Mutual Information
MICASE	The Michigan Corpus of Academic Spoken English



CHAPTER 1

INTRODUCTION

1.1 Background of the Study

English has emerged as the global lingua franca, largely due to the expansion of globalization (Anthony, 2018). As globalization connects people, cultures, technology, and economies across the globe, English plays a vital role in communication and interaction among people who do not share the same native language, including the Thai population. In Thailand, the awareness of the importance of the English language has influenced the rising number of International Programs, Bilingual Programs, English Programs (EP), and Mini English Programs (MEP) in both government and private schools over the past two decades (Keyurawong, 2010). This growth of English education requires stronger motivation and effort from students to overcome language learning obstacles. Furthermore, teachers play a pivotal role in providing guidance and facilitating student success in their language acquisition journey.

Vocabulary is a core component in foreign language teaching and learning. Students can communicate a broader range of meanings in various situations when they know and can use more words effectively (Coxhead, 2006). Mastering new vocabulary, especially synonyms, requires significant effort from students. Laufer (1990) observed that learning synonyms is further complicated by the fact that some terms may be interchangeable in certain situations but not in others; for example, *arguments* can be *strong* or *powerful*, but *tea* can be only *strong* and an *engine* only *powerful*. Teachers play a crucial role in addressing this complexity by employing suitable teaching techniques to help students overcome confusion and grasp subtle meanings.

This study investigated three synonymous adjectives: *beautiful*, *attractive*, and *gorgeous*. These words are among the top 3,000 frequently used words and appear as synonyms in the thesaurus section according to the Longman Dictionary of Contemporary English Online (LDOCE, online). These three words have similar meanings, but their subtle differences can confuse learners, making it difficult to choose the right one for a given situation.

Corpus tools are specialized computer software or online platforms designed to analyze extensive text collections, known as corpora. Widely used general corpora containing spoken and written texts include the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC), each containing over 100 million words from diverse sources. These tools provide valuable insights into language usage by enabling users to extract linguistic information, such as word frequency, collocations, and grammatical patterns. Integrating corpus-based approaches into course instruction can help students develop a deeper understanding of English usage.

1.2 Research Objectives

This research aimed to investigate the differences among three synonymous adjectives, *beautiful*, *attractive*, and *gorgeous*, across various genres using the COCA corpus. Additionally, the study examined the similarities and differences of the target synonyms in terms of meanings, formality, adverb and noun collocations, and semantic preferences.

1.3 Research Questions

- 1) How are the synonyms *beautiful*, *attractive*, and *gorgeous* distributed across different genres?
- 2) What are the similarities and differences among the near-synonymous adjectives *beautiful*, *attractive*, and *gorgeous* in terms of meanings, adverb and noun collocations in relation to semantic preferences?

1.4 Definitions of Key Terms

1) Corpus/Corpora: An electronic collection of spoken or written texts representing authentic language use in various contexts.

2) Corpus of Contemporary American English (COCA): A large general corpus containing over one billion words of text from eight genres: spoken, fiction, popular magazines, newspapers, academic texts, TV and movie subtitles, blogs, and other web pages.

3) Collocations: A combination of two or more lexical (not grammatical) words often found together or in close proximity, e.g., *dire straits and make sense* (Timmis, 2015).

4) Formality: The language uses style or the appropriate words for formal situations. For example, *utilize* might be chosen in professional writing or speech, while *use* would be more common in informal situations (LDOCE, online).

5) Genres: The different types or styles of texts represented in the Corpus of Contemporary American English (COCA), classified into eight categories: spoken, fiction, popular magazines, newspapers, academic texts, TV and movie subtitles, blogs, and other web pages.

6) Mutual Information (MI): A corpus statistic used to measure the strength of association between words. It is a statistical method frequently employed by corpus linguists to investigate collocations (Szudarski, 2017).

7) Synonymy: A relationship of “sameness of meaning” that may hold between two words (Jackson & Amvela, 2007).

8) Synonyms: A word that has the same or similar meaning to another word.

1.5 Scope of the Study

This analysis explored the differences between the three synonymous adjectives *beautiful*, *attractive*, and *gorgeous* across eight genres in the Corpus of Contemporary American English (COCA). The similarities and differences among the target synonyms, in terms of meanings, adverb and noun collocations, and semantic preferences, were analyzed using the Longman Dictionary of Contemporary English Online (LDOCE, online), Merriam-Webster Dictionary Online, and the COCA corpus. The top 30 noun and adverb collocates were selected based on their frequency and a Mutual Information (MI) score of three or higher (≥ 3).

1.6 Significance of the Study

A corpus-based study of near-synonyms is valuable for language learners and educators in several ways. Firstly, investigating the distribution of synonyms across eight genres helps clarify the domains in which the target synonyms are frequently used. This knowledge allows learners to choose appropriate words for specific situations.

Secondly, the study enhances the understanding of the overlapping meanings of the target synonyms, which dictionaries may only sometimes fully explain. Lastly, analyzing collocations that frequently co-occur with the target words helps illuminate their typical usage patterns and subtle distinctions in meaning.

1.7 Organization of the Study

This corpus-based study of the near-synonyms *beautiful*, *attractive*, and *gorgeous* consists of five chapters as follows:

- 1) Chapter One provides an introduction, including the study's background, research objectives, research questions, definitions of key terms, scope, significance, and organization.
- 2) Chapter Two reviews the relevant literature, encompassing the concept of corpora, synonymy, criteria for distinguishing synonyms, and previous studies.
- 3) Chapter Three explains the research methodology, including data collection procedures and analysis.
- 4) Chapter Four presents the results of the study.
- 5) Chapter Five provides a summary of the findings, discussion, conclusion, implications, limitations, and recommendations for future research.

CHAPTER 2

REVIEW OF LITERATURE

2.1 The Concept of Corpora

Corpora is a plural form of “corpus.” A corpus is a collection of used language naturally occurring under various circumstances (Brazil, 1995). However, a corpus is not just a collection of isolated words or sentences that occur naturally; it consists of spoken and/or written texts stored on a computer (Timmis, 2015). This enables researchers to use special software to conduct automatic searches and gain insights into the structure and regularity of naturally occurring language (Szudarski, 2017). Corpus tools offer several functions to support both quantitative and qualitative corpus analysis. For example, we can determine the frequency of words or phrases in quantitative corpus analysis and identify collocations and grammar patterns using concordance lines in qualitative studies. Numerous online corpora are accessible to researchers, offering valuable resources suited to their research purposes, such as the Corpus of Contemporary American English (COCA), the British National Corpus (BNC), and the Michigan Corpus of Academic Spoken English (MICASE). Corpora serve as valuable resources that provide new insights into language. Essentially, corpus studies reveal different language functions and underscore the central role of context in the analysis of linguistic behavior (Flowerdew, 2012).

2.1.1 The Corpus of Contemporary American English (COCA)

The present study employed the Corpus of Contemporary American English (COCA), a large corpus online database comprising over one billion words of text from eight genres: spoken, fiction, popular magazines, newspapers, academic texts, TV and movie subtitles, blogs, and other web pages. These diverse genres represent samples of American English language usage in various contexts. Importantly, COCA offers researchers electronic tools for conducting quantitative and qualitative research in linguistics, e.g., the analysis of semantic preference, identification of lexical patterns, and access to statistical tools for examining word frequency.

2.1.1.1 Genres in the Corpus of Contemporary American English (COCA)

According to the COCA corpus, genres refer to different categories or types of texts in the corpus database, which are evenly divided among eight genres as follows.

Spoken comprises transcripts of unscripted conversations collected from over 150 different TV and radio programs, such as Newshour (PBS), Good Morning America (ABC), Today Show (NBC), etc.

Fiction comprises short stories and plays collected from literary magazines, children's magazines, popular magazines, the first chapters of first-edition books from 1990 to the present, and movie scripts.

Magazines comprise nearly 100 different magazines, with a good mix (overall and by year) between specific domains (news, health, home and gardening, women, financial, religion, sports, etc.). A few examples are Time, Men's Health, Good Housekeeping, Cosmopolitan, Fortune, Christian Century, Sports Illustrated, etc.

Newspapers comprise ten newspapers from the US, including USA Today, New York Times, Atlanta Journal-Constitution, San Francisco Chronicle, etc. It provides a good mix of different sections of the newspaper, such as local news, opinion, sports, and financial.

Academic Journals comprise more than 200 different peer-reviewed journals covering the full range of academic disciplines, with a good balance among education, social sciences, history, humanities, law, medicine, philosophy/religion, science/technology, and business.

Web Pages comprise the web genres of academic, argument, fiction, info, instruction, legal, news, personal, promotion, and review web pages (by Serge Sharoff). They are taken from the US portion of the Corpus of Global Web-Based English (GloWbE).

Blogs comprise texts that Google has classified as blogs. These are further categorized into the web genres of academic, argument, fiction, info, instruction, legal, news, personal, promotion, and review web pages. The blogs were sourced from the US portion of the GloWbE corpus.

TV/Movie Subtitles comprise subtitles from OpenSubtitles.org and later the TV and movie corpora. Studies have shown that the language from these shows and movies is even more colloquial/ core than the actual spoken corpora data.

2.2 The Concept of Synonymy

In modern semantics, the term “synonymy” refers to a relationship of “sameness of meaning” that may hold between two words (Jackson & Amvela, 2007).

A synonym is defined as a word or phrase with the same meaning as another word or phrase (Webb & Nation, 2017). Synonyms can be classified into two types: strict and loose synonyms.

Strict synonyms (also known as absolute or perfect synonyms) are interchangeable in all possible contexts of use. This means one word can be freely substituted by its synonym without affecting the original meaning, style, or connotation (Cruse, 1986). Linguists argue that there are very few perfect synonyms or they do not exist at all.

Loose synonyms (also known as near-synonyms) are not interchangeable with one another in every context and may overlap in meanings to a greater or lesser degree (Jackson & Amvela, 2007). For example, *kid* and *child* share the same core meaning but differ in degree of formality. The *kid* is commonly used in everyday conversation, while the *child* is more prevalent in formal or professional contexts (LDOCE, online), as shown in the following sentences:

- My *kid* wants to go to the park after school.
- Every *child* deserves access to quality education.

2.3 Criteria for Distinguishing Synonyms

According to Jackson and Amvela (2007), words with similar meanings can be distinguished using several criteria. These include the specific dialect of English being

used, the level of formality appropriate for different contexts, the connotations associated with each word, the way words typically co-occur with others or collocations, and the categories of meaning to which they belong, or semantic preferences. These factors help in understanding the subtle differences between words that might initially appear to be interchangeable, allowing for more precise and contextually appropriate language use.

2.3.1 Senses of Meaning

A pair of synonyms may mostly share a denotation but may differ in connotation (Jackson & Amvela, 2007). Denotative or conceptual meaning refers to logical meaning that is considered the central factor in linguistic communication (Leech, 1981, P9). Similarly, the Cambridge Dictionary Online defines “denotation” as the primary meaning of a word, not including the feelings or ideas that people may connect with the word. Meanwhile, “connotative meaning is the communicative value of expression based on what it refers to, exceeds and above its pure conceptual content” (Leech, 1981). The following examples illustrate synonym pairs with distinct connotations (indicated in brackets):

childlike	childish (inappropriate behavior)
determine	stubborn (unwillingly)
house	home (with pleasant emotion)
slim	skinny (unattractive way)
unique	odd (unusual sense)

(Longman Dictionary of Contemporary English, online version)

ambiguous	equivocal (deliberately)
famous	notorious (disreputably)
new	novel (strikingly)
obtain	procure (with effort)
persuade	inveigle (with ingenuity or deceit)

(Jackson & Amvela, 2007)

2.3.2 Degree of Formality

Near-synonyms can be distinguished by the communication style between interlocutors in various circumstances. The set of words that share similar meanings may be used in a formal situation, while the other is used in less formal or informal situations or even belongs to slang or colloquial English (Jackson & Amvela, 2007). Near-synonyms must be carefully selected to maintain the flow of conversation and convey a sense of authority. Here are some synonym pairs that differ in terms of formality:

Formal	Informal
acquire	get
facilitate	help
inquire	ask
terminate	end
utilize	use

(Longman Dictionary of Contemporary English, online version)

2.3.3 Collocations

Firth (1957), a pioneer in the study of collocation, famously stated, “You shall know a word by the company it keeps.” This quote underscores the significance of understanding words through their context. In the Firthian sense, collocations are words that frequently and habitually appear together in a language. The Longman Dictionary of Contemporary Online defines “collocation” as the way in which some words are often used together or a particular combination of words used in this way. For instance, the word *stiff* often collocates with specific nouns, such as *stiff drink* and *stiff breeze*, while *heavy* commonly collocates with nouns like *heavy rain* and *heavy meal*. Collocations generally do not occur in isolation; they can be a combination of words that often co-occur. Collocations are not entirely predetermined or specific; instead, they are probabilistic events that arise from the repeated combination by language speakers (O’Keeffe et al., 2007). This concept shifts focus from individual words to pairs of words embedded in chunks of meaning, and usage and collocation are now

acknowledged as an aspect of vocabulary description and pedagogy (e.g., Lewis, 2000; McCarthy & O'Dell, 2017).

2.3.4 *Semantic Preference*

Stubbs (2001) states that semantic preference refers to “the relation, not between individual words, but between a lemma or word form and a set of semantically related words.” For example, *pay* often co-occurs with words in financial context: *pay rent*, *pay bills*, *pay taxes*, and words indicating actions: *pay attention*, *pay respect*. Therefore, semantic preference correlates with collocation and colligation principles but emphasizes groups of semantic categories instead of individual words (Baker, 2006). Stubbs (2001) noted that “even a category of semantic preference will be open-ended but will contain frequent and typical members.” Exploring semantic preference helps learners understand the relationship between semantic groups and how sets of semantically related words are used within contexts.

2.4 Previous Studies

Over the past two decades, there has been a notable increase in research on corpus-based analysis of near-synonyms. This field of study provides significant advantages for teachers and learners, enabling them to comprehend the nuances of language and select the most appropriate words confidently in any context.

Bergdahl (2009) examined three adjectives, *beautiful*, *handsome*, and *good-looking*, to determine if they are synonymous and to explore their similarities and differences. This investigation compared data from five contemporary dictionaries, one etymological dictionary, and the Corpus of Contemporary American English (COCA). The study revealed discrepancies between the data provided by dictionaries and the COCA corpus, with COCA offering more examples to clarify word definitions and usage contexts than the dictionaries. The researcher found that the three adjectives can be considered synonymous in specific contexts, such as when describing buildings, objects, and animals, but cautioned that connotative meanings should be considered when describing a person's appearance. The analysis also revealed that regarding gender aspects, *beautiful* is primarily used to describe women, *handsome* is more

frequently associated with men, and *good-looking* is considered gender-neutral based on the findings.

Ikonen (2022) investigated two synonymous adjectives, *big* and *large*, regarding meanings, frequency, and noun collocates. The definitions of both target words were collected from two well-known online dictionaries of contemporary English: the Collins English Dictionary and the Macmillan Dictionary. Both dictionaries defined these synonyms as referring to the physical size of people or things, and they are interchangeably used in these terms. The study demonstrated that *large* is typically used in a literal sense, while *big* can also have figurative meanings and is often used in idioms and slogans. The frequency results of the two words showed that *big* has twice as many occurrences as *large* in the Corpus of Contemporary American English (COCA). The study also focused on the immediate noun collocates. *Big* consistently co-occurs with nouns related to sports, e.g., *big leagues*, *big plays*; on the other hand, *large* more frequently co-occurs with nouns from scientific and medical terminology, e.g., *large proportion*, *large capacity*. Additionally, *big* and *large* can sometimes be used interchangeably with specific noun phrases, e.g., *big difference* or *large difference*, *big company* or *large company*.

Nugroho (2018) examined two synonyms, *rich* and *wealthy*, collecting data from the Merriam-Webster Online Dictionary and Thesaurus and the British National Corpus (BNC). The study data were analyzed using Sketch Engine (SkE) software. According to the definitions, *rich* and *wealthy* have similar core meanings but are not interchangeably used in certain contexts. The results from the BNC showed that *rich* is more commonly used than *wealthy* in spoken and written texts and has broader usage. It seems that *rich* typically collocates with nouns related to both humans and non-human beings, e.g., *rich peasant*, *rich country*, while *wealthy* generally describes humans, e.g., *wealthy merchants*, *wealthy industrialists*. Moreover, the study revealed the results in social aspects to describe human beings that *rich* and *wealthy* usually co-occur with words expressing the powerful and successful people in society.

Phoocharoensil (2022) conducted a corpus study to differentiate between three synonymous adjectives, *primary*, *main*, and *major*, by using the Longman Dictionary of Contemporary English (2014) and the data from the Corpus of Contemporary American English (COCA). The study focused on the distribution in different genres,

the meanings, and the collocation analysis based on the frequency and the Mutual Information (MI) score of three at the minimum or higher. The findings of the study showed that *primary*, *main*, and *major* shared some core meanings and are primarily used in formal contexts. The study categorized associated noun collocations of the target synonyms into semantic preferences to identify the relation between themes and each adjective. The results found that *concern* is often associated with all three synonymous adjectives, e.g., *primary concern*, *main concern*, and *major concern*. Moreover, certain adjectives are commonly associated with specific themes, e.g., *primary* with health and elections, *main* with places, food, or literature, and *major* with sports or business.

Thongpan (2022) investigated three synonyms, *far*, *distant*, and *remote*, regarding the degree of formality in their distribution across genres, collocations, and semantic preferences using the COCA corpora as a primary research instrument. The top 30 noun and verb collocates of the target synonyms were selected from the most frequent words with an MI score of 3 or higher (≥ 3) due to their strong association with collocation. The results showed that *far* is most commonly used in informal contexts, whereas *distant* and *remote* tend to be found in formal contexts. *Far* has the fewest collocations compared to its synonyms, which means it cannot be used with as many different nouns as *distant* and *remote*. The analysis of semantic preferences revealed that *far* and *remote* typically describe the distance between places, e.g., *far land*, *remote area*. Conversely, *distant* is more commonly used to express the distance of spatial elements, e.g., *distant galaxy*, *distant horizon*. The findings also indicated that all three synonyms had shared only two semantic categories: spatial elements and places. This suggests the potential for deeper semantic distinctions among these synonyms.

Previous corpus-based studies of near-synonyms have demonstrated effective techniques in distinguishing subtle meaning differences and identifying common collocation and grammatical patterns. Furthermore, the COCA corpus is a widely accepted tool for corpus-based researchers with its extensive and balanced collection of texts across various genres. The MI score function helps quantify the strength of the association between the target synonyms and their collocates, providing a reliable measure of their relationship. Additionally, near-synonyms may only sometimes be interchangeable in all situations. Analyzing semantic preferences helps identify these

limitations and restrictions, allowing us to gain a deeper understanding of the specific meanings that each synonym contributes to different contexts.



CHAPTER 3

RESEARCH METHODOLOGY

This chapter illustrates the approach and processes employed to analyze the target near-synonyms. It consists of three primary sections: (1) target near-synonyms, (2) data collection, and (3) data analysis.

3.1 Target Near-Synonyms

This research investigates three synonymous adjectives: *beautiful*, *attractive*, and *gorgeous*. While two major online dictionaries, the Longman Dictionary of Contemporary English Online (LDOCE, online) and Merriam-Webster Dictionary, list these words as synonyms in their thesaurus entries, the study delves deeper into exploring any subtle variations in their meaning and usage. This study aims to explore several aspects of the target words: their occurrences across eight different genres in resources provided by the Corpus of Contemporary American English (COCA), their conceptual meaning through comparison from two well-known online dictionaries, and their collocations with semantic preferences.

According to the Longman Dictionary of Contemporary English, recognized as a widely used online dictionary among English learners, the three target synonyms, *beautiful*, *attractive*, and *gorgeous*, are found in the Longman Communication 3000 most frequent word list, which represents the core common words using in the authentic communication in both speaking and writing. To illustrate, *beautiful*, *attractive*, and *gorgeous* are classified as S1, S2, and S3, indicating their ranks in the top 1000, 2000, and 3000 spoken words, respectively. Also, both *beautiful* and *attractive* are denoted by the symbol W2, signifying their ranks in the top 2000 written words, while *gorgeous*, being solely unmarked, suggests a more limited presence in written materials compared to its spoken usage. The analysis of the Longman Corpus Network has revealed that the 3,000 most common words in spoken and written English encompass 86% of the language, allowing learners to understand a vast majority of written material (LDOCE, 2009). Therefore, learning these fundamental words is highly influential, as it aids

students in cultivating proficient comprehension and communication abilities in the English language.

3.2 Data Collection

In this corpus-based study, *beautiful*, *attractive*, and *gorgeous* were investigated for their definitions, degree of formality, and collocations, drawing upon the two well-known online dictionaries: the Longman Dictionary of Contemporary English and the Merriam-Webster Dictionary. The selected dictionaries are significant resources for language learners and researchers as they offer comprehensive information, such as definitions in various aspects, synonyms, example sentences, thesaurus sections, grammar patterns, word frequency descriptions, and pronunciation audios. The rationale for using online dictionaries in this study is influenced by their convenient accessibility from any location with an internet connection and their wide range of regularly updated information, surpassing printed versions. Therefore, these two online dictionaries were used to reference the meanings of three target synonyms and compare the conceptual meanings of each word.

In addition to the online dictionaries, the investigation relied on COCA as the primary corpus resource. The latest version of the COCA corpus was released in March 2020. It is a widely used corpus among researchers, teachers, and learners due to its extensive database, which contains over a billion words of texts in various genres, with each genre balanced with approximately 20 million words per year from 1990 to 2019 (Davies, 2024). These genres include spoken language, fiction, popular magazines, newspapers, academic texts, TV and movie subtitles, blogs, and other web pages. COCA's extensive range of texts makes it a valuable resource for corpus-based analysis. Researchers can delve into various aspects such as word frequency, phrase usage, collocations, and grammatical structures across eight different genres, covering both formal and informal contexts. Therefore, it is reasonable to use the COCA corpus as the primary resource to facilitate the investigation of genres and collocation features in this study.

3.3 Data Analysis

This analysis began by examining the similarities and distinctions in the meaning among the three synonymous adjectives: *beautiful*, *attractive*, and *gorgeous*. The data was collected from two online dictionaries, the Longman Dictionary of Contemporary English and the Merriam-Webster Dictionary. The process involved utilizing the definitions and example sentences provided in the dictionaries to identify the aspects of meanings and collocation patterns of each word.

Next, COCA was employed to examine the frequency and distribution of the three synonyms across eight genres, intending to answer the first research question. Additionally, the data shed light on the typical usage of the target words, revealing whether they are predominantly used in formal or informal contexts.

After that, the top 30 noun and adverb collocates of *beautiful*, *attractive*, and *gorgeous* were investigated in order to identify the collocational patterns of the near-synonyms as well as the semantic preferences. The criteria for selecting noun and adverb collocates were based on their frequency and the Mutual Information (MI) value. The collocates were identified based on MI scores of at least three, indicating their significance in a strong collocational association (Schmitt, 2010; Cheng, 2012; Szudarski, 2017; Phoocharoensil, 2022). Then, the list of target noun and verb collocates were subsequently categorized based on their semantic preferences, and the collocates that shared some semantic preferences were grouped into the same theme.

CHAPTER 4

RESULTS AND DISCUSSION

This chapter delves into the study's findings in response to the research questions regarding the similarities and differences of meanings, distribution across various genres, the common noun and adverb collocations, and the semantic preferences of the three target adjectives: *beautiful*, *attractive*, and *gorgeous*.

4.1 Meanings

Two notable dictionaries were used to identify the meanings of the target words: the online version of the Longman Dictionary of Contemporary English and the Merriam-Webster Dictionary. Since the three target words are listed as synonyms in their thesaurus entries, this research aims to delve deeper to explore any subtle variations in their meaning, as illustrated in Table 1.

Table 1

Meanings of "Beautiful," "Attractive," and "Gorgeous" from the Longman Dictionary of Contemporary English Online (2024) and Merriam-Webster Dictionary Online (2024)

Synonyms	Definitions	
	Longman Dictionary of Contemporary English Online (2024)	Merriam-Webster Dictionary Online (2024)
<i>beautiful</i>	1. someone or something that is beautiful is extremely attractive to look at <i>e.g., She was even more beautiful than I had remembered.</i> 2. very good or giving you great pleasure <i>e.g., The weather was beautiful.</i>	1. having qualities of beauty <i>e.g., You have the most beautiful smile.</i> 2. generally pleasing <i>e.g., What a beautiful day for a picnic!</i>
<i>attractive</i>	1. someone who is attractive is good looking, especially in a way that makes you sexually interested in them <i>e.g., Women seem to find him attractive.</i> 2. pleasant to look at	1. arousing interest or pleasure <i>e.g., An attractive smile</i> 2. having or relating to the power to attract <i>e.g., Attractive forces between molecules</i>

Synonyms	Definitions	
	Longman Dictionary of Contemporary English Online (2024)	Merriam-Webster Dictionary Online (2024)
	<p><i>e.g., Kitchen utensils should be attractive as well as functional.</i></p> <p>3.having qualities that make you want to accept something or be involved in it <i>e.g., I must say, it's a very attractive offer.</i></p>	
gorgeous	<p>1.extremely beautiful or attractive <i>e.g., You look gorgeous, Maria.</i></p> <p>2.extremely pleasant or enjoyable <i>e.g., The hotel room had a gorgeous view.</i></p>	<p>1.splendidly or showily brilliant or magnificent <i>e.g., Your baby is absolutely gorgeous! Sunset in Hawaii are just gorgeous.</i></p>

Table 1 reveals that beautiful, attractive, and gorgeous seem to have a similar core meaning. They all convey generally positive descriptions of beauty and pleasantness. These descriptions can be applied to a broad range of subjects, encompassing people, places, and objects, or even evoke emotional admiration for intangible concepts. However, a deeper analysis uncovers subtle distinctions between these near-synonymous adjectives. Notably, *beautiful* emerges as the most general term, simply signifying something aesthetically pleasing and adhering to established beauty standards.

In contrast, *attractive* and *gorgeous* contain more subtle implications, relating to the degree of intensity and extending beyond physical attributes. *Gorgeous* carries a significantly stronger connotation of intensity compared to *beautiful*. As demonstrated in Table 1, both dictionaries define *gorgeous* by using adverb modifiers such as *extremely beautiful* or *extremely pleasant*. Moreover, *attractive* encompasses a broader range of positive qualities and the power to arouse interest. This versatility allows it to be used for things appealing beyond the physical, making it applicable in many different contexts.

In conclusion, a comparative analysis of definitions from two online dictionaries, the Longman Dictionary of Contemporary English and the Merriam-Webster Dictionary reveals a similar core meaning of “pleasing to look at” among these target words. However, they are not interchangeable in every context due to the subtle

distinctions in intensity and connotations related to physical and non-physical features that arouse interest. Therefore, these three target words are considered near-synonyms. Speakers should carefully select the most appropriate term for the situation to convey a specific message effectively.

4.2 Frequency and Distribution of the Synonyms Across Genres

In response to the first research question, the results regarding the overall frequency of the three target synonyms are presented first, followed by the distribution across eight genres with which the synonyms are commonly associated.

Table 2

The Overall Frequency of “Beautiful,” “Attractive,” and “Gorgeous” Across Eight Genres From COCA

Genre	<i>beautiful</i>		<i>attractive</i>		<i>gorgeous</i>	
	Frequency	Per million	Frequency	Per million	Frequency	Per million
Spoken	15,350	121.69	1,959	15.53	2,078	16.47
Fiction	21,734	183.69	3,097	26.17	2,088	17.65
Magazines	12,671	100.49	4,302	34.12	2,022	16.04
Newspapers	7,502	61.62	2,792	22.93	1,051	8.63
Academic texts	3,827	31.95	3,068	25.61	125	1.04
TV/movie subtitles	32,306	252.24	2,227	17.39	3,090	24.13
Blogs	13,590	105.67	3,377	26.26	1,984	15.43
Webpages	13,788	110.97	2,759	22.20	1,611	12.97
Total	120,768		23,581		14,049	

Table 2 clearly illustrates that *beautiful* occurs with the highest frequency. With a total of 120,768 tokens, *beautiful* appears over five times more often than *attractive* (23,581 tokens), and *gorgeous* (14,049 tokens) appears the least among the other two

synonyms. This substantial difference in frequency suggests that *beautiful* is a more general term for describing aesthetics and applies to a broader range of things. *Attractive* and *gorgeous*, on the other hand, have lower frequencies in every genre compared to *beautiful*, suggesting that they might be used less and selectively in specific contexts.

Table 3 shows the distribution patterns of three near-synonyms, *beautiful*, *attractive*, and *gorgeous*, across eight distinct genres from COCA. The two synonyms, *beautiful* and *gorgeous*, are similar in their distribution across genres, as seen in the top three. *Beautiful* and *gorgeous* share the same ranks in the following genres: TV/movie subtitles (252.24 and 24.13 per million), fiction (183.69 and 17.65 per million), and spoken (121.69 and 16.47 per million). This shared dominance suggests a possible link between these words that are more common in spoken text and informal communication.

Table 3

Distribution of “Beautiful,” “Attractive,” and “Gorgeous” According to Frequency

Genre	<i>beautiful</i>		Genre	<i>attractive</i>		Genre	<i>gorgeous</i>	
	Frequency	Per million		Frequency	Per million		Frequency	Per million
TV and movie subtitles	32,306	252.24	Magazines	4,302	34.12	TV and movie subtitles	3,090	24.13
Fiction	21,734	183.69	Blogs	3,377	26.26	Fiction	2,088	17.65
Spoken	15,350	121.69	Fiction	3,097	26.17	Spoken	2,078	16.47
Webpages	13,788	110.97	Academic texts	3,068	25.61	Magazines	2,022	16.04
Blogs	13,590	105.67	Newspapers	2,792	22.93	Blogs	1,984	15.43
Magazines	12,671	100.49	Webpages	2,759	22.20	Webpages	1,611	12.97
Newspapers	7,502	61.62	TV and movie subtitles	2,227	17.39	Newspapers	1,051	8.63
Academic texts	3,827	31.95	Spoken	1,959	15.53	Academic texts	125	1.04
Total	120,768		Total	23,581		Total	14,049	

In contrast, *attractive* shows a high frequency in written texts, ranking highly in the top five genres: magazines (34.12 per million), blogs (26.26 per million), fiction (26.26 per million), academic texts (25.61 per million), and newspapers (22.93 per million). Unlike *beautiful* and *gorgeous*, which are primarily found in spoken language, *attractive* commonly appears across various written genres. This suggests a higher formality associated with *attractive*, which could also be adapted to formal and informal written styles depending on the context.

4.3 Collocations

Responding to the second research question, noun and adverb collocates of the three synonymous adjectives, *beautiful*, *attractive*, and *gorgeous*, were examined and analyzed. The top 30 noun and adverb collocates were selected from COCA based on their frequency and an MI value of 3 or higher, indicating a significant statistic value between co-occurring words with a strong relationship (Szudarski, 2017). Furthermore, semantic preferences of these noun and adverb collocates are identified and discussed in the following sections.

4.3.1 Noun Collocations

The top-30 noun collocates of *beautiful*, *attractive*, and *gorgeous* from COCA are illustrated in Table 4.

Table 4

Top-30 Frequency List of Noun Collocates of “Beautiful,” “Attractive,” and “Gorgeous” in COCA

	<i>beautiful</i>			<i>attractive</i>			<i>gorgeous</i>		
	Noun collocates	Frequency	MI Value	Noun collocates	Frequency	MI Value	Noun collocates	Frequency	MI Value
1	woman (3661)	5273	5.47	woman (909)	1464	5.78	woman (200)	345	4.34
	women (1612)		3.74	women (555)		4.40	women (145)		3.40
2	girl (1871)	2384	5.14	man (270)	475	3.21	day	215	3.09
	girls (513)		3.94	men (205)		3.32			

	<i>beautiful</i>			<i>attractive</i>			<i>gorgeous</i>		
	Noun collocates	Frequency	MI Value	Noun collocates	Frequency	MI Value	Noun collocates	Frequency	MI Value
3	day	2149	3.44	option (216)	276	6.27	girl (145)	205	4.49
				options (60)		4.40	girls (60)		3.98
4	thing (1604)	1388	3.35	girl (160)	250	3.92	color (74)	118	4.10
	things (784)		4.10	girls (90)		3.69	colors (44)		5.15
5	place (962)	1252	3.33	alternative (150)	184	5.83	hair	114	4.75
	places (290)		3.37	alternatives (34)		5.50			
6	eyes	831	3.66	place (128)	160	3.28	views (53)	103	4.84
				places (32)		3.66	view (50)		3.22
7	wife	754	3.95	features (83)	152	4.53	eyes	96	3.66
				feature (69)		4.93			
8	baby	692	3.82	target (87)	142	4.72	beach (40)	77	4.15
				targets (55)		5.61	beaches (37)		7.11
9	children (433)	678	3.76	price (72)	136	3.36	dress	60	5.20
	child (245)		3.94	prices (64)		4.03			
10	face	557	3.25	candidate (87)	127	4.65	wife	54	3.25
				candidates (40)		3.82			
11	hair	520	3.83	quality (66)	98	3.48	guy	54	3.17
				qualities (32)		5.26			
12	daughter	511	3.91	female (48)	93	3.36	face	52	3.50
				females (45)		5.29			
13	boy	484	3.36	investment	87	4.74	scenery	44	8.25
14	lady	472	4.61	offer (50)	80	3.02	man	49	3.03
						3.16			

<i>beautiful</i>			<i>attractive</i>			<i>gorgeous</i>			
Noun collocates	Frequency	MI Value	Noun collocates	Frequency	MI Value	Noun collocates	Frequency	MI Value	
offers (30)									
15	voice	459	3.42	guy	67	3.22	photos	42	4.42
16	music	432	3.19	lady	64	4.13	piece	40	3.48
17	city	431	4.02	force	58	3.16	body	36	3.34
18	song	418	4.06	person	57	3.84	voice	35	3.13
19	morning	401	3.02	design	51	3.06	ladies	33	4.72
20	country	367	3.73	opportunities	50	4.14	display	33	4.50
21	home	367	3.28	destination	49	5.99	creature	32	6.00
22	dress	363	4.72	proposition	48	6.04	place	32	3.06
23	piece	360	3.56	package	45	4.66	screen	32	3.58
24	house	335	3.35	market	41	3.16	flowers	30	4.67
25	flowers	326	5.09	choice	39	3.04	setting	28	3.85
26	mind	291	3.18	idea	38	3.46	smile	28	3.77
27	night	281	3.38	nuisance	37	7.60	weather	27	3.76
28	story	269	3.36	males	33	4.85	garden	26	3.79
29	skin	263	3.56	prospect	31	5.02	models	26	3.84
30	pictures	257	3.55	appearance	30	3.84	skin	26	3.17

Table 4 presents the top-30 noun collocations that frequently co-occur with the near-synonymous adjectives *beautiful*, *attractive*, and *gorgeous*. As can be seen, all three adjectives share some noun collocates. *Woman* is the most frequent noun, co-occurring with all three adjectives, followed by *girl*, *lady*, and *place*. This observation suggests a strong association between the target adjectives and common noun collocates related to descriptions of females and settings.

While the majority of shared noun collocates among the three target synonyms are related to females, interestingly, two noun collocates associated with males, *man*, and *guy*, are found in the shared noun collocates of *attractive* and *gorgeous*. Notably, *man* is the second most frequent word in the top-30 list for *attractive* after *woman*, whereas *beautiful* has only one noun collocate referring to the male gender, *boy*. This suggests that *attractive* and *gorgeous* are more frequently associated with nouns referring to men, particularly adult males when compared to *beautiful*.

Additionally, *attractive* demonstrates a greater variety of noun collocates beyond those of *beautiful* and *gorgeous*, such as *option*, *feature*, *target*, *candidate*, and

investment. This suggests that *attractive* can be used not only to express aesthetic appeal but also to convey desirability or value in a broader sense, including in business fields.

Moreover, *beautiful* and *gorgeous* share approximately one-third of their noun collocates, including *day, eyes, wife, hair, flowers, face, dress, piece, skin, and voice*, which indicates shared usage across various contexts. The considerable number of overlapping noun collocates between *beautiful* and *gorgeous* suggests similarities in their usage, particularly when describing physical appearance, natural phenomena, and tangible objects. This observation highlights their common association with aesthetic appeal and admiration.

In conclusion, Table 1 illustrates that the three adjectives share some similar core meanings. Additionally, some noun collocates are associated with expressing physical beauty and eliciting admiration for humans, places, and objects. However, it is crucial to consider the subtle distinctions in intensity degree and connotation of meaning for the target words *beautiful*, *attractive*, and *gorgeous*. This consideration helps determine which adjective is most appropriate for the specific situation.

The next step in the research procedure involved analyzing the semantic preferences of the three target synonyms to group their noun collocates based on similarities in meaning. This analysis will reveal the unique semantic field inhabited by the target synonyms. The results of the semantic preferences of noun collocates of *beautiful* are presented in Table 5.

Table 5

Semantic Preference of Noun Collocates of “Beautiful.”

1. HUMAN	woman, girl, wife, baby, daughter, boy, lady, child
2. HUMAN ATTRIBUTES	eyes, face, hair, skin, voice
3. SETTING	city, country, home, house, place
4. ARTISTIC CREATION	song, music, pictures, story
5. TIME PERIOD	day, morning, night
6. NON-SPECIFIC OBJECT	thing, piece
7. MENTALITY	mind
8. CLOTHING	dress
9. LIVING THING	flower

Table 5 illustrates the semantic preferences of noun collocates for the adjective *beautiful*. The table categorizes these collocates into nine thematic groups, offering insights into the diverse contexts in which *beautiful* is commonly employed. The first theme, HUMAN, contains *woman, girl, wife, baby, daughter, boy, lady, and child*, as exemplified in (1), and the second, HUMAN ATTRIBUTES, contains *eyes, face, hair, skin, and voice*, as can be seen in (2). The third theme, SETTING, includes noun collocates that can be applied to surroundings, e.g., *city, country, home, house, and place*, as shown in (3). The fourth theme, including *song, music, pictures, and story*, as exemplified in (4), is ARTISTIC CREATION. The fifth theme, TIME PERIOD, indicates the periods of the day, which contain *day, morning, and night*, as in (5). Two nouns, *thing*, and *piece* have been assigned to NON-SPECIFIC OBJECT, as in (6). Only one noun collocate is found in each of the MENTALITY, CLOTHING, and LIVING THING categories, which are *mind, dress, and flower*, as shown in (7-9), respectively.

(01) She was the most ***beautiful woman*** I had ever met.

(02) You have such ***beautiful hair***. Look at those eyes. Wow!

(03) I live in a ***beautiful house*** with a view of the mountains.

(04) This incredibly ***beautiful song*** with deep meaning can touch our souls.

(05) Adam and I had spent a ***beautiful day*** on a slice of heaven known as Alleppey in Kerala, India.

(06) It was the most ***beautiful thing*** I have ever seen.

(07) ... she is an inspired writer who can convey the thoughts of her ***beautiful mind*** while not leaving the listener feeling left behind.

(08) I found this ***beautiful dress*** in a random closet.

(09) A garden full of ***beautiful flowers***.

The analysis reveals that most noun collocates of *beautiful* are frequently associated with human subjects and general environments, with the top four semantic domains being HUMAN, HUMAN ATTRIBUTES, SETTING, and ARTISTIC CREATION. This suggests that *beautiful* is primarily used to express admiration for people, the surrounding environment, and artistic works. Additionally, the prevalence of semantic areas highlights a strong association of the adjective with human beings,

particularly women and girls, along with human attributes. This may indicate a potential tendency for the adjective to be used more frequently in describing the physical appearance or beauty of females.

Table 6

Semantic Preference of Noun Collocates of “Attractive.”

1. HUMAN	woman, man, girl, guy, lady, female, males, person, candidate
2. PRODUCT ATTRIBUTES	feature, price, quality, design, package, appearance
3. BUSINESS ACTIVITY	offer, proposition, investment
4. CHOICE	choice, option, alternative
5. CHANCE	opportunities, prospect
6. AIM	target, destination
7. ASPECT	nuisance, idea
8. SETTING	market, place
9. POWER	force

The analysis of nouns frequently used with the word *attractive*, as depicted in Table 6, reveals nine different themes illustrating the adjective's wide range of applications. These themes illustrate the wide range of noun collocates in which this adjective is utilized. While two themes, HUMAN and SETTING, overlap with those found for *beautiful* and *gorgeous*, seven themes emerge that indicate the distinctive usage patterns. The first theme, the HUMAN category, pertains to the appeal of individuals across gender identities and societal roles, encompassing terms such as *women, men, girls, guys, and candidates*, as illustrated in (10). The second theme, PRODUCT ATTRIBUTES, contains words related to the descriptions of products, including *feature, price, quality, design, package, and appearance*, as shown in (11). The third theme, BUSINESS ACTIVITY, includes *offer, proposition, and investment*, as exemplified in (12). The fourth theme, CHOICE, contains *choice, option, and alternative*, as in (13). The fifth theme, CHANCE, encompasses *opportunities* and *prospect*, as shown in (14). Two noun collocates are found in the AIM theme: *target* and *destination*, as in (15). The seventh theme, ASPECT, contains two noun collocates, *nuisance* and *idea*, as shown in (16), and another two noun collocates, *market* and *place*,

which are in the SETTING theme, as exemplified in (17). The last theme is POWER, which contains only one noun collocate, *force*, as shown in (18).

(10) She was an *attractive woman* in her early 40s.

(11) It was another nice meal at the very *attractive price* of just \$6.85 a la carte.

(12) The changes made real estate a very *attractive investment*, and demand soared.

(13) Homeschooling will become an increasingly *attractive option* for these parents.

(14) Getting their hemoglobin levels up to normal seemed an *attractive prospect*, an obvious way of helping.

(15) His growing fame apparently made him an *attractive target* to prosecutors.

(16) An abandoned house is an *attractive nuisance*. It attracts kids with rocks.

(17) They really want to make the Far East an *attractive place* for people to live.

(18) In the most basic definition, gravity is the *attractive force* between two objects.

Interestingly, the semantic preference themes for noun collocates of *attractive* overlap with those of *beautiful* in only two categories: HUMAN and SETTING. The remaining seven themes primarily relate to business management concepts, such as PRODUCT ATTRIBUTES, BUSINESS ACTIVITY, CHOICE, CHANCE, and AIM. This suggests that while *attractive* can describe people we find appealing, both females and males, it extends beyond physical beauty to encompass the business field.

Table 7

Semantic Preference of Noun Collocates of “Gorgeous.”

1. HUMAN	woman, girl, wife, ladies, man, guy, models
2. HUMAN ATTRIBUTES	body, eyes, face, hair, skin, smile, voice
3. SETTING	beach, garden, setting, scenery, view, place
4. VISUAL REPRESENTATION	color, photos, display, screen
5. LIVING THING	creature, flowers
6. CLOTHING	dress
7. TIME PERIOD	day
8. WEATHER	weather

9. NON-SPECIFIC OBJECT	piece
------------------------	-------

Table 7 presents the semantic preferences of noun collocates for the adjective *gorgeous*, organizing them into nine thematic categories. Most themes overlap with *beautiful*, except for VISUAL REPRESENTATION and WEATHER. HUMAN ATTRIBUTES and SETTING are the only two themes that overlap with *attractive*. For noun collocates of *gorgeous* in the HUMAN theme, the list encompasses *woman, girl, wife, ladies, man, guy, and models*, as shown in (19). Followed by HUMAN ATTRIBUTES, which contains *body, eyes, face, hair, skin, smile, and voice*, as in (20). The third theme, SETTING, includes *beach, garden, setting, scenery, weather, view, and place*, as exemplified in (21). The next theme is VISUAL REPRESENTATION, which includes *color, photos, display, and screen*, as in (22). The fifth theme, LIVING THING, includes *creature and flowers*, as shown in (23). Only one noun collocate is found in each of the CLOTHING, TIME PERIOD, WEATHER, and NON-SPECIFIC OBJECT themes, which are *dress, day, weather, and piece*, as exemplified in (24-27), respectively.

(19) When I started my job, I instantly fell for this **gorgeous woman**.

(20) I would love her to play up her **gorgeous eyes** with some shadow and liner.

(21) Diners can enjoy the **gorgeous setting** in the air-conditioned veranda dining room or on a seaside patio.

(22) The texture differs a bit, and the incredibly **gorgeous color** is enhanced.

(23) You can use an elegant vase to hold **gorgeous flowers** like a single stem of a white lily or an orchid.

(24) ...at everyone, her eyes full of tears, her **gorgeous dress** sweeping and rustling.

(25) Wow. It is a really **gorgeous day**, huh? Would you like me to take you out somewhere?

(26) Hey, everybody, we are looking at **gorgeous weather** coming up just in time for July 4th.

(27) Look at that **gorgeous piece** of bread. It is like a big baked cloud.

It is clear that the common noun collocates associated with *gorgeous* often fall into themes similar to the semantic preferences of *beautiful*, such as HUMAN, HUMAN ATTRIBUTES, and SETTING. However, the unique presence of the VISUAL REPRESENTATION and WEATHER themes of *gorgeous* suggests a stronger association with describing visual aesthetics or the beauty of the natural environment. This is evident in the distinct nouns collocating within the overlapping themes, such as *beach*, *garden*, *creature*, and *weather*. This implies that while *gorgeous* and *beautiful* share similar usage patterns, *gorgeous* tends to emphasize visually aesthetic qualities of nature. Also, *gorgeous* is used to describe both males and females, similar to *attractive*, as evidenced by the shared collocates of *woman*, *girl*, *man*, and *guys* in the HUMAN theme.

4.3.2 Adverb Collocations

Following the analysis of common noun collocates for *beautiful*, *attractive*, and *gorgeous* using the COCA corpus, the research extended to examine adverbs frequently used alongside these adjectives.

Table 8

Top-30 Frequency List of Adverb Collocates of “Beautiful,” “Attractive,” and “Gorgeous” in COCA

Rank	<i>beautiful</i>			<i>attractive</i>			<i>gorgeous</i>		
	Adverb collocates	Frequency	MI Value	Adverb collocates	Frequency	MI Value	Adverb collocates	Frequency	MI Value
1	most	5884	5.60	more	2498	5.52	absolutely	301	7.75
2	so	4905	3.80	very	1715	6.12	so	264	3.52
3	very	2062	4.03	most	704	4.89	most	248	4.14
4	more	1848	3.55	less	657	6.35	how	113	3.05
5	how	1610	3.84	so	452	3.55	really	77	3.43
6	really	1026	3.23	particularly	216	6.45	drop-dead	45	4.15
7	absolutely	457	5.25	physically	166	8.30	simply	39	4.05
8	quite	293	3.73	especially	141	5.00	incredibly	27	6.37
9	incredibly	213	6.24	quite	125	4.86	stunningly	27	10.86
10	truly	172	4.38	really	121	3.31	perfectly	17	4.98

Rank	<i>beautiful</i>			<i>attractive</i>			<i>gorgeous</i>		
	Adverb collocates	Frequency	MI Value	Adverb collocates	Frequency	MI Value	Adverb collocates	Frequency	MI Value
11	stunningly	162	10.34	extremely	117	6.57	unbelievably	15	8.86
12	strikingly	104	8.55	sexually	112	8.28	totally	15	3.98
13	hauntingly	100	11.17	increasingly	83	6.28	quite	14	3.16
14	breathtakingly	86	10.76	least	76	3.31	impossibly	11	8.53
15	extremely	77	3.61	highly	69	5.40	truly	11	3.52
16	extraordinarily	75	6.99	conventionally	57	10.76	insanely	10	8.84
17	particularly	72	3.30	incredibly	56	6.67	strikingly	9	8.12
18	spectacularly	57	8.23	economically	53	7.78	ridiculously	9	7.77
19	naturally	55	4.06	equally	52	6.14	utterly	9	5.80
20	strangely	54	6.15	reasonably	52	7.58	flat-out	7	8.51
21	amazingly	50	6.41	pretty	42	3.77	downright	6	6.64
22	perfectly	48	3.38	rather	40	3.72	effortlessly	6	7.76
23	equally	45	3.58	fairly	31	5.16	equally	6	3.77
24	unbelievably	43	7.28	politically	30	5.76	extremely	6	3.03
25	exceptionally	43	6.38	visually	28	7.38	amazingly	5	6.19
26	achingly	42	9.91	strikingly	25	8.85	positively	5	5.06
27	impossibly	40	7.29	financially	21	6.40	naturally	5	3.70
28	astoundingly	37	7.91	terribly	20	6.22	achingly	4	9.62
29	starkly	32	7.94	extraordinarily	19	7.37	breathtakingly	4	9.44
30	exquisitely	32	7.70	exceptionally	18	7.48	visually	4	5.32

Table 8 presents the top-30 adverb collocates of three synonymous adjectives, *beautiful*, *attractive*, and *gorgeous*, which were selected based on their frequency and MI values of 3 or higher (≥ 3) in the COCA. As can be seen, all three adjectives share some common adverbs, including *most*, *so*, *really*, *quite*, *incredibly*, *extremely*, and *equally*. It can be observed that *beautiful* and *gorgeous* share a considerable number of adverbs related to both the degree of aesthetic appreciation and emotional intensity towards beauty. In contrast, *attractive* exhibits a wider variety of adverb collocates, including those not commonly associated with *beautiful* and *gorgeous*: *physically*, *sexually*, *increasingly*, *economically*, and *financially*. These findings correlate with the results from the noun collocation analysis. This clear evidence indicates that *attractive* has a broader range of usage, encompassing physical and non-physical aesthetic expressions, as well as economic and financial descriptions.

To determine the similarities and differences of adverb collocates more precisely, the top 30 adverbs that co-occur with the three synonymous adjectives were grouped into semantic themes based on semantic preferences, as illustrated in Table 9-11.

Table 9

Semantic Preference of Adverb Collocates of “Beautiful.”

1. LEVEL/DEGREE	most, more, very, so, really, quite, absolutely, extremely, perfectly, equally, incredibly, extraordinarily, spectacularly, exceptionally, strikingly
2. EMOTION	stunningly, breathtakingly, amazingly, unbelievably, achingly, impossibly, astonishingly, hauntingly, starkly
3. QUALITY	exquisitely, truly, naturally, strangely
4. EMPHASIS	particularly
5. EXCLAMATION	how

Table 9 provides an analysis of the semantic preferences of adverb collocates for the adjective *beautiful*. The table categorizes these collocates into five thematic groups. The first theme, LEVEL/DEGREE, encompasses adverbs expressing varying degrees of intensity, illustrating the spectrum of beauty from moderate to extreme, e.g., *quite*, *very*, *spectacularly*, as shown in (28), (29), and (30). The second theme, EMOTION, reveals how adverbs can capture the robust emotional response that beauty often evokes, including positive and ambiguous aspects, such as *stunningly*, *achingly*, and *hauntingly*, as illustrated in (31), (32), and (33). The third theme, QUALITY, focuses on adverbs describing the nature or style of the beauty itself, e.g., *strangely* and *naturally*, as in (34) and (35). The last two themes, EMPHASIS, and EXCLAMATION, each contain only one adverb, *particularly* as shown in (36) and *how* as shown in (37).

(28) The music here is skillfully written and often ***quite beautiful***, but it has never drawn me in as much as works by other...

(29) They have a ***very beautiful*** property right on the lake with the most horrendous stairs imaginable!

(30) He designed a ***spectacularly beautiful*** home for the Bellingham Herald.

(31) Yosemite and Zion National Parks, both stunningly beautiful valleys, would be gated communities, access available only for a few wealthy...

(32) ..., a run of unsentimental small towns and achingly beautiful sunsets on a scenic loop between Fredericksburg and Williamsburg.

(33) A seventh-grade girl starts humming a hauntingly beautiful tune from “Phantom of the Opera.”

(34) ...One visit convinced me that no traveler in Turkey should miss this unique and strangely beautiful region.

(35) ...instead, say hello to sunscreen and sunblock and better protect your naturally beautiful skin from the harmful rays of the sun.

(36) What happened to that particularly beautiful scarf you were wearing at the fashion show?

(37) ...fruits are also a big part of the art. Look at how beautiful it is on an apple.

As can be seen, the majority of adverb collocations with *beautiful* fall under the first and second themes, which involve adverbs indicating varying degrees of intensity and emotional response. This evidence suggests that speakers typically use adverbs associated with *beautiful* to express different levels of feeling and emotion in response to the beauty they perceive. This aligns with the understanding that perceptions and reactions can vary greatly among individuals.

Table 10

Semantic Preference of Adverb Collocates of “Attractive.”

1. LEVEL/DEGREE	more, very, most, less, so, least, extremely, equally, incredibly, extraordinarily, exceptionally, increasingly, strikingly, terribly, fairly, really, highly, quite, pretty, rather
2. ACADEMIC	economically, financially, politically
3. APPEARANCE	physically, visually, sexually
4. EMPHASIS	particularly, especially
5. QUALITY	conventionally, reasonably

Table 10 presents the semantic preferences of adverb collocates for the adjective *attractive*, organized into five thematic categories. Most adverb collocates fall under the first theme, LEVEL/DEGREE, which contains a variety of intensifiers that commonly co-occur with *attractive*, e.g., *less*, *fairly*, and *most*, as demonstrated in (38), (39), and (40). The second theme, ACADEMIC, includes adverbs related to academic or intellectual contexts, such as *economically*, *financially*, and *politically*, as exemplified in (41). The third theme, APPEARANCE, comprises adverbs that specify physical and visual aspects, e.g., *physically*, *visually*, and *sexually*, as in (42). The fourth theme, EMPHASIS, includes *particularly* and *especially*, highlights the specificity of certain aspects, as shown in (43). Lastly, the QUALITY theme contains *conventionally* and *reasonably*, as in (44).

(38) ...Second, recent retirees are increasingly likely to accept employment in **less attractive** positions but hope this will lead...

(39) He was a tall, muscular, **fairly attractive** man, and I remember thinking, “Why not?” By the time...

(40) At the same time, they are not exactly the **most attractive** job candidates in the market due to age.

(41) Such strategies include placing a strong emphasis on attaining the full, **economically attractive** improvement in end use energy efficiency, both for existing and new buildings; ...

(42) The confident, active, **physically attractive** individualism of their stars is a vital ingredient in their selling of consumer products.

(43) These private equity funds have several features that make them **particularly attractive** for an empirical study.

(44) ...for men is a lot broader. What is a **conventionally attractive** man? He can be buff or thin or have a few extra pounds.

Evidently, *attractive* shares a significant number of LEVEL/DEGREE collocates with *beautiful*, but lacks representation in the EMOTION category. This suggests that *attractive* is often modified by adverbs emphasizing degree rather than emotional impact, aligning with its higher prevalence in formal contexts compared to *beautiful*. Thus, *attractive* features unique adverb collocates related to physical

appearance and academic contexts, indicating a broader semantic scope. These findings underscore the subtle yet significant differences in the connotative meanings and usage patterns of *attractive*, reflecting its versatile role in communication.

Table 11

Semantic Preference of Adverb Collocates of “Gorgeous”

1. LEVEL/DEGREE	most, so, really, quite, equally, extremely, incredibly, strikingly, absolutely, perfectly, totally, utterly, simply, drop-dead, flat-out
2. EMOTION	amazingly, breathtakingly, achingly, insanely, ridiculously, stunningly, unbelievably, impossibly, downright,
3. QUALITY	effortlessly, truly, naturally, positively, simply,
4. APPEARANCE	visually
5. EXCLAMATION	how

Table 11 demonstrates an analysis of the semantic preferences of adverb collocates for the adjective *gorgeous*, divided into five thematic groups. The primary group, LEVEL/DEGREE, is a common theme shared by all three adjectives with some overlap of adverb collocates. The adverbs commonly co-occur with *gorgeous* in this theme, e.g., *absolutely*, *strikingly*, and *totally*, as demonstrated in (45). The second theme, EMOTION, also overlaps with *beautiful*. These adverbs evoke diverse emotional responses to the experience of beauty in various contexts. The adverbs associated with *gorgeous* in this group, e.g., *amazingly*, *breathtakingly*, and *unbelievably*, as shown in (46). QUALITY, the third theme, encompasses *effortlessly*, *truly*, *naturally*, and *positively*, expressing the target adjective's inherent nature, as in (47). The fourth theme, APPEARANCE, consists of one adverb, *visually*, which relates to the visual aspect or perception, as exemplified in (48). Finally, one adverb *how* has been assigned to the EXCLAMATION theme, as in (49).

(45) The restaurant is *absolutely gorgeous*, and the staff goes above and beyond to make your experience the best ...

(46) I blushed. Hot MAX, our *unbelievably gorgeous* customer, came through the espresso line every morning at nine.

(47) ...When a woman is naturally gorgeous, she really can be distracting with too much makeup. Less is really...

(48) The film is visually gorgeous.

(49) Look how gorgeous she looks in her little dress.

The analysis reveals a significant overlap in the semantic domains of LEVEL/DEGREE, EMOTION, and QUALITY between *gorgeous* and *beautiful*, suggesting similar usage of adverb collocates. Both adjectives predominantly focus on expressing heightened intensity and emotional responses related to beauty. However, *gorgeous* demonstrates a greater propensity for adverbs conveying intensity and strong emotional responses, as evidenced by collocates such as *drop-dead*, *flat-out*, *downright*, *insanely*, and *ridiculously*. These adverbs, absent in the *beautiful* adverb collocates list, suggest that *gorgeous* is often used in more informal and spoken contexts. This aligns with the classification of *gorgeous* in the Longman Dictionary of Contemporary English, online version, which labels it as among the top 3,000 spoken words (S3) and identifies it as an informal language. Additionally, only *gorgeous* and *attractive* share adverb collocates in the APPEARANCE theme, suggesting a stronger emphasis on visual appearance and the power of attractiveness in their expressions.

CHAPTER 5

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

This chapter consists of six primary sections: (1) Summary of the findings, (2) Discussion, (3) Conclusion, (4) Implications, (5) Limitations, and (6) Recommendations.

5.1 Summary of the Findings

This corpus-based study aimed to investigate three near-synonyms, *beautiful*, *attractive*, and *gorgeous*, using the Corpus of Contemporary American English (COCA) and two online dictionaries. The following section summarizes the findings regarding the distribution across eight genres, similarities and differences in meaning, noun and adverb collocations, and semantic preferences.

5.1.1 Meanings

The research examined the definitions of *beautiful*, *attractive*, and *gorgeous* using two online dictionaries: the Longman Dictionary of Contemporary English and Merriam-Webster Dictionary. The investigation revealed that these words share a similar core meaning of “pleasing to look at.” However, they cannot be used interchangeably in all contexts due to subtle differences in intensity and connotation. Given the shared core meaning and the slight variations, it is most appropriate to classify the three target words as near-synonyms.

5.1.2 Genres Distribution of the Target Words

The Corpus of Contemporary American English (COCA) was employed to investigate the distributional patterns of the target words. The analysis revealed a significant difference in the frequency of *beautiful* among its synonyms, with *beautiful* appearing most frequently, followed by *attractive* and then *gorgeous*. Interestingly, *beautiful* and *gorgeous* share the same top-3 genres: TV/movie subtitles (252.24 tokens), fiction (183.69 tokens), and spoken (121.69 tokens). Notably, they are found

least frequently in academic texts. On the other hand, *attractive* often occurs in written texts, as its top-5 genres are magazines, blogs, fiction, academic texts, and newspapers.

5.1.3 Noun Collocations

The investigation of noun collocates and their semantic preferences for the three target synonyms indicates that *woman* appears as the most frequent noun collocate for all three target adjectives: *beautiful*, *attractive*, and *gorgeous*, followed by *girl*, *lady*, and *place*, which are common to all adjectives. *Beautiful* and *gorgeous* share a considerable number of nouns collocates, while *attractive* encompasses a broader range of nouns, particularly those related to the business field.

There are two common semantic preference themes that are shared by all three adjectives: HUMAN and SETTING. *Beautiful* and *gorgeous* mostly share similar themes, whereas *attractive* exhibits a broader variety correlated with its noun collocates, particularly in the business domain, including PRODUCT ATTRIBUTES, BUSINESS ACTIVITY, CHOICE, CHANCE, AIM, and ASPECT.

5.1.4 Adverb Collocations

The study revealed that *beautiful*, *attractive*, and *gorgeous* can be intensified using similar adverbs, e.g., *most*, *so*, *really*, *quite*, *incredibly*, *extremely*, and *equally*, which are categorized under the LEVEL/DEGREE theme. However, *beautiful* and *gorgeous* evoke more than just intensity; they focus on describing a strong emotional response to aesthetics, falling under the EMOTION theme, while *attractive* primarily emphasizes just the level of intensity yet lacks strong emotional connotations. Additionally, *attractive* can be modified with a broader range of adverbs, showcasing its versatility in usage, including those related to APPEARANCE and ACADEMIC, e.g., *physically*, *sexually*, *economically*, and *financially*.

5.2 Discussion

The results of this study are discussed in terms of similarities and differences in meaning, distribution across genres, formality degree, collocations, and semantic preferences.

5.2.1 Meanings

This study investigated the semantic similarities and differences between *beautiful*, *attractive*, and *gorgeous*, which appeared synonymous in two online dictionaries' thesaurus entries. It was found that all three target adjectives share a similar core denotation but differ in their intensity degree and connotative meanings, which indicate their status as loose synonyms or near-synonyms. This observation aligns with Jackson and Amvela's (2007) and Stubbs's (2001) notion that synonyms may have the same denotation or conceptual meaning but differ in connotation, which may carry different associative or emotive meanings.

As suggested by Stubbs (2001), different dictionaries can vary considerably in distinguishing between a word's literal meaning and its suggested or implied meanings. The results show that *beautiful* carries a broader definition of being pleasing, extending beyond simple physical attractiveness to encompass various qualities of beauty. It is evident that *beautiful* commonly co-occurs with concrete and abstract nouns, e.g., *beautiful music* or *a beautiful bunch of flowers*. Next, *gorgeous* is defined as exceptional physical beauty or attractiveness, representing a higher intensity level than its synonyms. The study found that *gorgeous* frequently co-occurs with concrete nouns, e.g., *gorgeous view* or *gorgeous cake*. While *gorgeous* and *beautiful* share similar core meanings, they differ in intensity, with *gorgeous* conveying a greater degree of beauty. In contrast, *attractive* has a more complex meaning, defined as the ability to draw positive attention, encompassing not only physical features but also personality traits or any other quality that can capture someone's interest. This broader scope is reflected in its more comprehensive range of usage compared to the other two words. The sample sentences demonstrated that *attractive* is often associated with noun collocates representing concrete and abstract qualities, e.g., *attractive young woman* or *attractive offer*.

Similar to previous studies (e.g., Bergdahl (2009); Nugroho (2018); Petchrat & Phoocharoensil, 2017; Phoocharoensil (2021); Phoocharoensil (2022); Ikonen (2022)) that identified limitations in the information provided by dictionaries, this study also sought a more comprehensive understanding of word usages in real-world contexts by utilizing the COCA corpus. By analyzing corpus data, educators can gain insights

beyond dictionary definitions, leading to more effective teaching strategies and approaches that better reflect the complexities of natural language.

5.2.2 Genres Distribution of the Target Words

According to an analysis of the distribution of *beautiful*, *attractive*, and *gorgeous* across eight genres in COCA, the findings reveal that *beautiful* is the most frequent term, occurring in approximately 120,768 tokens. This is about five times more than *attractive* (23,581 tokens) and about eight times more than *gorgeous* (14,049 tokens). This indicates that *beautiful* is a more common term, potentially used more often in conversation than *attractive* and *gorgeous*.

Notably, *beautiful* and *gorgeous* dominate their top three most frequent genres: TV/movie subtitles, fiction, and spoken language, and they also display their lowest frequency in the exact domains of academic texts and newspapers. This observation implies that they are more commonly used in spoken language, which is a characteristic of informal communication. In contrast, *attractive* exhibits a high frequency in written texts, ranking highly in the top five genres: magazines, blogs, fiction, academic texts, and newspapers, which suggests a higher level of formality associated with *attractive*. This finding aligns with previous corpus-based studies, which suggest that certain near-synonyms can appear in different contexts with varying degrees of formality, e.g., *far*, *distant*, and *remote* (Thongpan, 2022), *error*, *fault*, and *mistake* (Phoocharoensil, 2020), *appropriate*, *proper*, and *suitable* (Petchrat & Phoocharoensil, 2017), and *problem* and *trouble* (Jirananthiporn, 2018).

5.2.3 Noun Collocations

This corpus-based research utilized COCA to gather the top-30 most frequent noun and adverb collocations for the adjectives *beautiful*, *attractive*, and *gorgeous* based on MI value greater than or equal to 3, which indicates consistency in the association between two words (Hunston, 2002; Cheng, 2012). The study demonstrates that *woman* emerged as the most frequent noun collocates of all three target synonyms, followed by *girl*, *lady*, and *place*, which were common noun collocates for all three. These findings suggest a strong association between the target adjectives and nouns related to humans, particularly females. While the terms *man* and *guy* were presented

as noun collocates for *attractive* and *gorgeous*, with *man* being the second most frequent noun collocate for *attractive*. This observation indicates that *attractive* and *gorgeous* are associated with noun collocates referring to both males and females, whereas *beautiful* is primarily associated with terms referencing females. Similar to a previous study by Bergdahl (2009), which found that the adjectives *beautiful*, *handsome*, and *good-looking* have gender-specific connotations, i.e., *beautiful* is more commonly used to describe women, while *handsome* is typically used to describe men.

As observed from the evidence regarding the semantic preference of noun collocates, all three adjectives share the same themes of HUMAN and SETTING, suggesting a potential for synonymy among all three target words. For the HUMAN theme, the individual collocates of all target words frequently appear with words referring to humans, such as *woman*, *girl*, *wife*, *daughter*, *lady*, *males*, and *guy*. However, the individual noun collocates associated with SETTING themes tend to differ more significantly across the three adjectives, e.g., *beautiful city*, *attractive place*, *gorgeous beach*.

Moreover, *gorgeous* and *beautiful* share a significant number of noun and adverb collocations, overlapping in almost all semantic fields except VISUAL REPRESENTATION and WEATHER. It can be inferred that *gorgeous* is often used to describe something tangible or visual beauty, e.g., *gorgeous display*, *gorgeous photos*, while *beautiful* has a broader sense of beauty that can extend to describe intangible qualities, e.g., *beautiful song*, *beautiful story*, *beautiful mind*.

Furthermore, apart from the two shared semantic fields, *attractive* extends its usage to encompass broader themes beyond those typically associated with the other two adjectives, e.g., PRODUCT ATTRIBUTES, BUSINESS ACTIVITY, CHOICE, CHANCE, and AIM. These broader uses for *attractive* cover the area of the business field with association to both concrete and abstract nouns, e.g., *attractive candidate*, *attractive option*, and *attractive prospect*.

5.2.4 Adverb Collocations

The findings in the adverb collocate section revealed that *beautiful*, *attractive*, and *gorgeous* share a significant number of adverbs in the LEVEL/DEGREE theme, e.g., *most*, *so*, *really*, *quite*, *incredibly*, *extremely*, and *equally*. Most of the shared

adverb collocates are considered weak collocates because they are commonly used with almost all adjectives in English (Petchrat & Phoocharoensil, 2017). Although QUALITY is another shared theme, the specific adverbs used with each adjective show little overlap. This suggests that while the adjectives share some similarities in terms of conveying intensity or degree, they also have distinctive ways of describing qualities.

Beautiful and *gorgeous* share a considerable number of adverbs related to the degree of aesthetic appreciation and emotional intensity towards beauty, falling under the LEVEL/DEGREE, EMOTION theme. This aligns with their prevalence in spoken language, which often conveys personal feelings and subjective evaluations. In contrast, *attractive* is primarily modified by adverbs of degree, with less emphasis on emotional intensity, including a wider range of adverb collocates in the ACADEMIC theme, e.g., *economically*, *financially*, and *politically*. This evidence is consistent with the results in genre analysis and noun collocation, which demonstrated that *attractive* is often found in written contexts and frequently co-occurs with business terms, suggesting a higher degree of formality and less emotive expression. The different adverb collocates and semantic preferences can indicate the loose synonyms or near-synonyms status of the three target adjectives, as this is one of the criteria used to differentiate near-synonyms (Phoocharoensil, 2010).

Furthermore, the adverb collocates in the APPEARANCE theme, e.g., *physically*, *visually*, and *sexually*, tend to occur more with *attractive* than its synonyms. This suggests that *attractive* carries a stronger connotation of physical allure and outward appearance. In contrast, *beautiful* and *gorgeous* encompass a broader range of qualities, including inner beauty, and tend to emphasize the degree of intensity and emotional impact associated with beauty.

This significant overlap in noun and adverb collocates and semantic preferences suggests that these three adjectives share similar usage patterns in some contexts; however, subtle differences in connotative and additional meanings exist. Due to these similarities and differences, the three adjectives cannot be used interchangeably in all contexts, which lends support to a number of studies (e.g., Petchrat & Phoocharoensil, 2017; Phoocharoensil & Kanokpermpoon, 2021; Lertcharoenwanich & Phoocharoensil, 2022; Chaengchenkit, 2023; Sittironnarit et al., 2022).

5.3 Conclusion

This corpus-based study aimed to explore two research questions. First, the study examined similarities and differences among *beautiful*, *attractive*, and *gorgeous* in terms of their meanings and distribution across genres. Second, the study investigated the typical noun and adverb collocates in relation to semantic preference. Drawing on evidence from two online dictionaries, the analysis revealed that *beautiful*, *attractive*, and *gorgeous* share a similar core meaning of “pleasing to look at” but differ in intensity and connotation; therefore, they cannot be interchangeably used in all contexts. The results from COCA demonstrated that *beautiful*, and *gorgeous* appear more frequently in spoken language, which is characteristic of informal communication. In contrast, *attractive* primarily appears in written genres and also potentially occurs in academic texts and newspapers, which indicates a higher degree of formality than the other two words. Finally, in terms of noun and adverb collocates, all three target words share some overlap. *Beautiful* and *gorgeous* exhibit a particularly strong connection, sharing a significant number of noun and adverb collocates in similar fields. In contrast, *attractive* has a more distinct set of collocates, encompassing broader semantic preferences beyond those typically associated with *beautiful* and *gorgeous*. For example, *attractive* collocates might include terms from the business field, e.g., *attractive price* and *attractive prospect*.

5.4 Implications

The findings of this study have significant implications for English language teaching. They suggest that incorporating COCA into the curriculum can enhance the student learning experience. A “learning by doing” approach, where students actively explore how native speakers use synonyms in real-world situations, can be particularly effective. For instance, instructors can guide students to search COCA for sentences containing synonyms like “hard” and “strong.” By analyzing the context surrounding each word, students uncover subtle differences in meaning and usage. This hands-on approach fosters a deeper understanding of the nuances of synonyms, including denotation, connotation, intensity, and formality.

COCA data provides concordance lines and collocate information that teachers can use to create tailored lessons and activities. Gap-filling exercises, for instance,

encourage students to search COCA and select the most suitable word or synonym based on context clues. This interactive approach promotes a more engaging and effective learning environment for English as a Foreign Language (EFL) students.

While students have generally relied heavily on thesauruses, which simply list synonyms without providing context, teachers should guide them in selecting the most appropriate word for a specific situation. Introducing students to up-to-date resources, such as reliable online dictionaries used in conjunction with English corpora like COCA, can further enhance their ability to discern subtle distinctions between synonyms. By encouraging students to consult both thesauruses and online dictionaries, teachers can empower them to verify definitions and connotations, leading to more precise and effective language use.

5.5 Limitations

This corpus-based study has several limitations. The restriction of noun and adverb collocates to the top-30 collocation lists means that other potential associations with the three target adjectives may have been overlooked due to their low frequency and Mutual Information (MI) scores in COCA. Notably, *beautiful* is among the top 1000 spoken words in the Longman Communication 3000-word list and has the highest frequency among its synonyms. This suggests that the study might not fully capture useful collocations for real-life conversations. For instance, *beautiful world* (1,147 times), *beautiful people* (1,137 times), and *beautiful life* (573 times) were not selected due to MI values below the study's threshold of 3, which indicates inconsistent association. Similarly, expanding the extraction to the top-40 lists would have included *beautiful smile* (253 times), *attractive people* (21 times), and *gorgeous spring* (21 times). Additionally, the adverb *how* frequently co-occurs with *beautiful* and *gorgeous* but is absent for *attractive*, as its occurrence would be in the top 10 with a frequency of 165 times, yet its MI value does not meet the study's criteria.

While the MI score is one of the most common statistical measurements in collocation analysis (Szudarski, 2018), it has limitations. Some collocations may show a strong statistical association but occur infrequently within the corpus, suggesting they may not be the most representative examples (Cheng, 2012). Thus, both frequency and

MI values should be considered and utilized collaboratively for more comprehensive and effective results, rather than relying exclusively on one measure.

5.6 Recommendations

This study primarily focused on examining the distribution, meaning, collocation, and semantic preferences using two online dictionaries alongside the Corpus of Contemporary American English (COCA) as the main analytical tool. COCA encompasses a comprehensive collection of written and spoken American English texts. However, it is important to consider other prominent English-speaking corpora, such as the British National Corpus (BNC), which comprises a diverse range of written and spoken texts in British English, providing authentic examples of usage. Additionally, for researchers interested in specialized studies, the Michigan Corpus of Academic Spoken English (MICASE) offers a valuable resource specifically for spoken language in academic settings. This corpus comprises over 1.8 million words collected from diverse academic contexts such as lectures, dissertation defenses, meetings, and service encounters and is openly accessible to the public. By incorporating additional corpora into future research efforts, researchers can achieve greater clarity in comparing and contrasting different varieties of English (such as American English and British English) or delve deeper into specific registers or dialects. This would allow for a more comprehensive and detailed understanding of the language, thereby enriching the depth and scope of their studies.

REFERENCES

- Anthony, L. (2018). *Introducing English for specific purposes*. Routledge.
- Bergdahl, M. (2009). *Is he beautiful, handsome or good-looking?: A study of putative synonymy in three adjectives* [Master's thesis, Luleå University of Technology, Sweden].
- Keyurawong, S. (2010). *Learning through English: Policies, challenges, and prospects: Insights from East Asia*. British Council.
- Chaengchenkit, R. (2023). A corpus-based study of the synonyms "Cease," "Halt," and "Stop." *LEARN Journal: Language Education and Acquisition Research Network*, 16(1), 473-494.
- Cheng, W. (2011). *Exploring corpus linguistics: Language in action*. Routledge.
- Coxhead, A. (2006). *Essentials of teaching academic vocabulary*. Houghton Mifflin Co.
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge University Press.
- Davies, M. (2024). *The corpus of contemporary American English*.
- Firth, J.R. (1957). *A synopsis of linguistic theory, 1930-1955. Studies in linguistic analysis*. Blackwell.
- Flowerdew, L. (2011). *Corpora and language education*. Palgrave Macmillan.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
- Ikonen, N. (2022). *Big and large: A corpus study on near-synonyms* [Bachelor's thesis, Tampere University].
- Jackson, H., & Zé Amvela, E. (2007). *Words, meaning and vocabulary: An introduction to modern English lexicology* (2nd ed.). Continuum.
- Jiranthiporn, S. (2018). Is this problem giving you trouble? A corpus-based examination of the differences between the nouns problem and trouble. *Thoughts*, 2(1), 1-25.

- Laufer, B. (1990). Words you know: How they affect the words you learn. *Further insights into contrastive linguistics*, 573-593.
- Leech, G. (1981). *Semantics: The study of meaning* (2nd ed). Penguin Books.
- Lertcharoenwanich, P., & Phoocharoensil, S. (2022). A corpus-based study of the near-synonyms: purpose, goal, and objective. *REFlections*, 29(1), 148-168.
- Lewis, M. (Ed.). (2000). *Teaching collocation: further developments in the lexical approach*. Oxford University Press.
- Longman Dictionary of Contemporary English*. (2009). Pearson Education.
- Longman Dictionary of Contemporary English Online. <https://www.ldoceonline.com>
- McCarthy, M., & O'Dell, F. (2017). *English collocations in use: How words work together for fluent and natural English: Self-study and classroom use*. Cambridge University Press.
- Merriam-Webster dictionary. <https://www.merriam-webster.com>
- Nugroho, D. Y. (2018). A corpus linguistics investigation of two near-synonymous words: Rich and wealthy. *English Education: Journal of English Teaching and Research*, 3(2), 118-127.
- O'keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.
- 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.
- Phoocharoensil, S. (2010). A corpus-based study of English synonyms. *International Journal of Arts and Sciences*, 3(10), 227-245.
- Phoocharoensil, S. (2020). Collocational patterns of the near-synonyms error, fault, and mistake. *The International Journal of Communication and Linguistic Studies*, 19(1), 1-17. <https://doi.org/10.18848/2327-7882/CGP/v19i01/1-17>

- Phoocharoensil, S. (2021). Semantic prosody and collocation: A corpus study of the near-synonyms persist and persevere. *Eurasian journal of applied linguistics*, 7(1), 240-258. <https://doi.org/10.32601/ejal.911269>
- 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), 968-975. <https://doi.org/10.34044/j.kjss.2021.42.4.34>
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.
- Sittironnarit, S., Khunasathitchai, K., Kosashunhanan, K., & Kumdee, S. (2022). A corpus-based analysis of English synonyms: acquire and obtain. *Journal of Roi Kaensarn Academi*, 7(9), 242-252.
- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Blackwell.
- Szudarski, P. (2017). *Corpus linguistics for vocabulary: A guide for research*. Routledge.
- Thongpan, N. (2022). A corpus-based study of English synonyms of the adjectives 'Far,' 'Distant,' and 'Remote.' *Journal of Positive School Psychology*, 3986-4001.
- Timmis, I. (2015). *Corpus linguistics for ELT: Research and practice*. Routledge.
- Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.