



**A CORPUS-BASED ANALYSIS OF SCENT AND
FRAGRANCE**

BY

THANAPHORN KINAMANEE

**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 2022**

**A CORPUS-BASED ANALYSIS OF SCENT AND
FRAGRANCE**

BY

THANAPHORN KINAMANEE

**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 2022**

THAMMASAT UNIVERSITY
LANGUAGE INSTITUTE

AN INDEPENDENT STUDY

BY

THANAPHORN KINAMANEE

ENTITLED

A CORPUS-BASED ANALYSIS OF SCENT AND FRAGRANCE

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

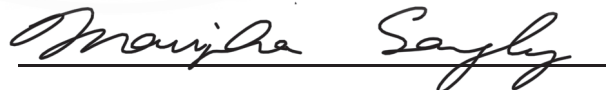
on June 20, 2023

Chairman



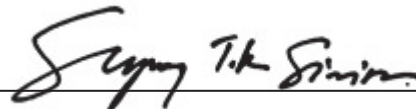
(Assistant Professor Vimalchaya Yanasugondha, Ph.D.)

Member and Advisor



(Assistant Professor Monnipha Somphong, Ph.D.)

Director



(Associate Professor Supong Tangkiengsiririn, Ph.D.)

Independent Study Title	A CORPUS-BASED ANALYSIS OF SCENT AND FRAGRANCE
Author	Thanaphorn Kinamanee
Degree	Master of Arts
Major Field/Faculty/University	English Language Teaching Language Institute Thammasat University
Independent Study Advisor	Assistant Professor Monnipha Somphong, Ph.D.
Academic Year	2022

ABSTRACT

This study aims to examine the similarities and differences between the noun synonyms *scent* and *fragrance* in terms of meanings, degrees of formality, and adjective collocations. To achieve this, two online dictionaries, namely the Longman Dictionary of Contemporary English (LDCE) and the Oxford Advanced Learners' Dictionary (OALD), along with the Corpus of Contemporary American English (COCA), were consulted. The findings reveal that *scent* and *fragrance* can be considered near-synonyms due to their shared core meanings. However, they are not interchangeable in every context, as *scent* has an additional definition related to *animal or human smell*, which is not present in *fragrance*. In terms of formality, both synonyms tend to be more commonly used in informal contexts, as evidenced by their prevalence in genres such as fiction and magazines. Analyzing the top 20 frequent adjective collocates of these synonyms, eight collocates are found to be shared, including *floral*, *fresh*, *heady*, *heavy*, *pleasant*, *spicy*, *sweet*, and *warm*. The study also highlights the benefits of both dictionaries and corpora in English vocabulary learning and teaching. Dictionaries serve as a primary tool for accessing brief word information, while corpora provide more comprehensive details that aid in distinguishing synonyms.

Keywords: synonyms, near synonyms, corpus-based analysis, concordance lines

ACKNOWLEDGEMENTS

Firstly, I am honored to begin by expressing my appreciation to my adviser, Assistant Professor Dr. Monipha Somphong, for her useful guidance and support during the independent study, without which the study could not have been accomplished.

My sincere appreciation should be also shown to Assistant Professor Dr. Vimolchaya Yanasugondha, the other Independent Study committee member, who made helpful suggestions for significant improvements during this study.

I am very also thankful to all the faculty members at the Language Institute, Thammasat University for their teaching which provided me with both theoretical knowledge and practical experience necessary to conduct this research.

Lastly, I would also like to thank my ELT friends, and staff members at the Language Institute, for their help and encouragement as well as my beloved family whose support enabled me to successfully complete this study.

Thanaphorn Kinamanee

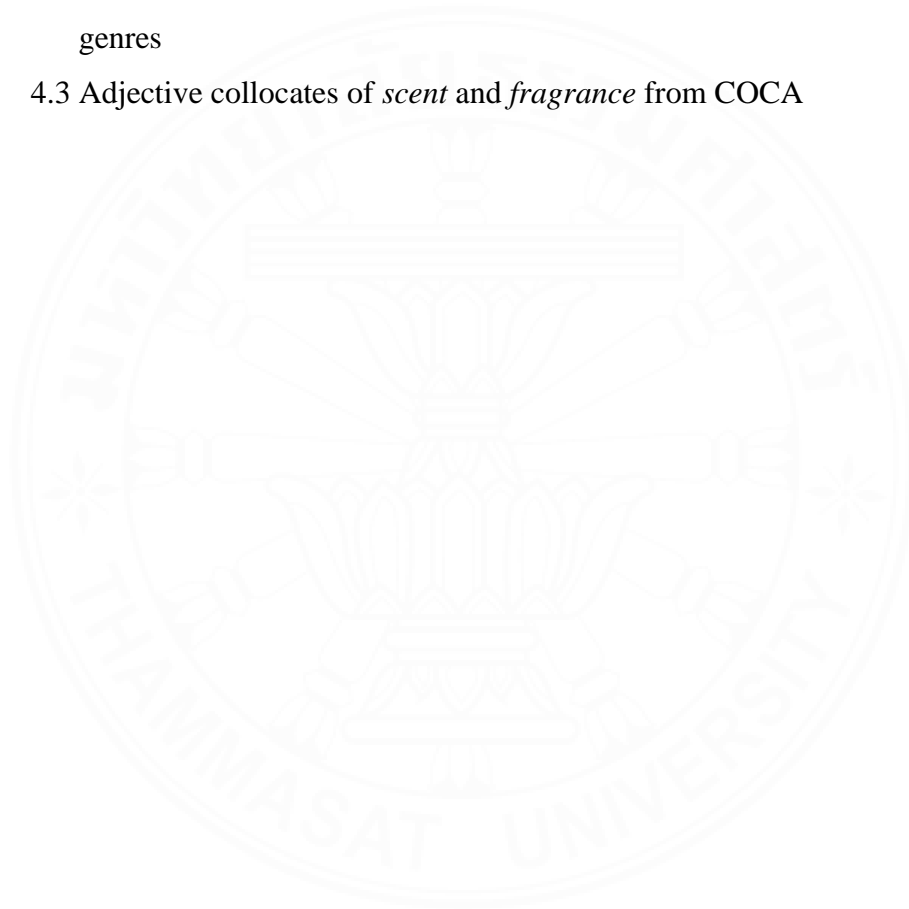
TABLE OF CONTENTS

	Page
ABSTRACT	(1)
ACKNOWLEDGEMENTS	(2)
LIST OF TABLES	(5)
LIST OF ABBREVIATIONS	(6)
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	5
2.1 Concept of synonyms	5
2.1.1 Absolute synonyms	5
2.1.2 Near synonyms	6
2.2 Corpus linguistics and language pedagogy	6
2.3 Criteria for distinguishing synonyms	7
2.3.1 Collocations	7
2.3.2 Degrees of formality	8
2.4 English language corpora	9

	(4)
2.5 Previous related studies	10
CHAPTER 3 RESEARCH METHODOLOGY	13
3.1 Selected synonyms	13
3.2 Research instruments	13
3.3 Data analysis	14
CHAPTER 4 RESULTS	15
4.1 Meanings	15
4.2 Degrees of formality	16
4.3 Adjective collocations	17
CHAPTER 5 DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS	19
5.1 Summary of the findings	19
5.2 Discussion	20
5.3 Conclusion	21
5.4 Implications	22
5.5 Recommendations for future research	23
REFERENCES	24

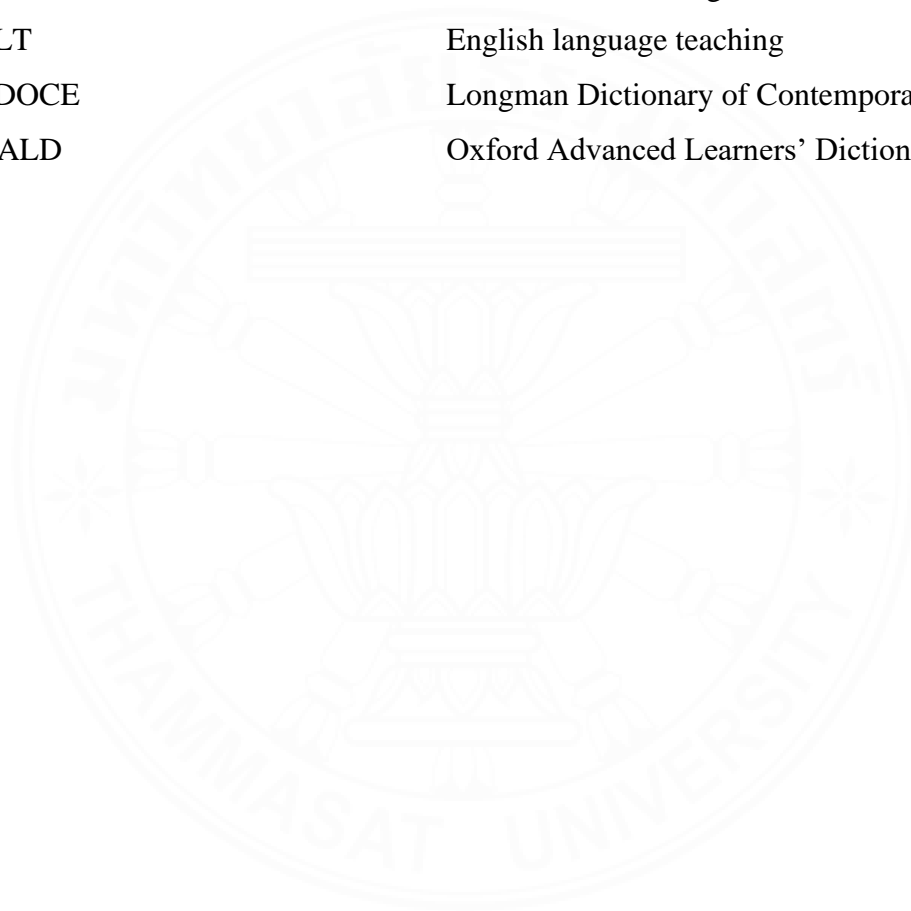
LIST OF TABLES

Tables	Page
4.1 Meanings of <i>scent</i> and <i>fragrance</i> from LDOCE (2018) and OALD (2023) in the online versions	15
4.2 Total frequency and division of <i>scent</i> and <i>fragrance</i> through eight genres	16
4.3 Adjective collocates of <i>scent</i> and <i>fragrance</i> from COCA	17



LIST OF ABBREVIATIONS

Abbreviations	Terms
BNC	British National Corpus
COCA	Corpus of Contemporary American English
DDL	Data-driven Learning
ELT	English language teaching
LDOCE	Longman Dictionary of Contemporary English
OALD	Oxford Advanced Learners' Dictionary



CHAPTER 1

INTRODUCTION

1.1 Background of the Study

Vocabulary knowledge is widely recognized as a crucial aspect of English language learning and instruction, as it supports the development of the four major language skills. However, mastering English vocabulary can pose challenges due to the sheer volume of words to be learned (Jackson and Amvela, 2007). Language learners often face difficulties when it comes to synonyms, which are words with similar meanings. The interchangeable usage of some synonyms within the same context can lead to confusion and further complicate the process of vocabulary acquisition.

Many previous research studies have shown that there are no English synonyms that are completely interchangeable in every context of use since they differ in several aspects, consisting of degrees of formality (e.g., Ruenroeng, 2014; Jirananthiporn, 2018), connotations (e.g., Partington, 1998; Stubbs, 1995), collocations (e.g., Gu, 2017; Jarunwaraphan and Mallikamas, 2020), semantic prosody (e.g., Liu and Espino, 2012; Uba, 2015; Phoocharoensil, 2021), and grammatical patterns (e.g., Petcharat and Phoocharoensil, 2017; Thongpan, 2022).

The synonyms *scent* and *fragrance* will be investigated in this study regarding their meanings, degrees of formality, and adjective collocations. The study was conducted with the purpose of distinguishing the target noun synonyms using the data derived from a massive American English corpus, the Corpus of Contemporary American English (COCA). The meanings and usages of both words were compared and contrasted by consulting two learner dictionaries, i.e., Longman Dictionary of Contemporary English (LDOCE) and Oxford Advanced Learners' Dictionary (OALD) before COCA data was employed to obtain more detailed information on the ways the target synonyms occur in authentic English.

1.2 Research Objectives

1) To examine the synonyms *scent* and *fragrance* in terms of their similarities and differences with regard to meanings, degrees of formality, and adjective collocations.

2) To compare and contrast the data obtained from the selected dictionaries with those obtained from the corpus.

1.3 Research Questions

1) What are the similarities and differences between the two synonymous nouns *scent* and *fragrance* regarding the meanings, degrees of formality, and adjective collocations?

2) What are the similarities and differences of the two nouns *scent* and *fragrance* between those found in the dictionaries and in the corpus data?

1.4 Definitions of Key Terms

The key terms definitions in the present study clarified below:

1) *Absolute synonyms* are words that can substitute for each other in every context without changing the messages' meaning, style, or connotation (Jackson & Amvela, 2000).

2) *COCA* stands for Corpus of Contemporary American English, which is produced by Mark Davies, considered as the largest corpus of American English with free online access. COCA is widely used by about a hundred thousand users, e.g., teachers, linguists, and researchers.

3) *Corpus* (*Corpora* for plural) is a computer-stored collection of both spoken and written texts (O'Keeffe et al., 2007), readily providing data for qualitative and quantitative analysis (Kayaoglu, 2013).

4) *Collocation* is a combination of words that occurs more regularly than would happen by coincidence in a language (Oxford Learner's Dictionary of Academic English, 2014, p. 132).

5) *Concordance lines* are a collection that exemplifies the use of a keyword in context that is produced by native English users collected in COCA.

6) *Genre* is a distinct form or style of art, film, literature, or music which is distinguished by its special features (Oxford Advanced Learner's Dictionary, 2015).

7) *Near-synonyms* are words that are considered as almost synonyms but cannot be substituted for each other in all exact same contexts. They may differ in their shades of denotations or connotations, collocations, or grammatical patterns (Inkpen & Hirst, 2006).

8) *Native English speakers* are people who speak English as their first language.

9) *MI score* stands for Mutual Information score which is a statistical association measure used to quantify the strength of the attraction between words as evidenced by their semantic co-occurrence in a text (Brezina & Gablasova, 2018).

10) *Synonym* is a word or expression in the same language that has the same or nearly the same meaning as another (Oxford Learner's Dictionary of Academic English, 2014, p. 811).

1.5 Scope of the Study

The two synonymous nouns *scent* and *fragrance* are the main focus of this study. The sources of data were online version of two learner dictionaries, namely Longman Dictionary of Contemporary English (LDOCE) and Oxford Advanced Learners' Dictionary (OALD), as well as a massive corpus representing American English, namely the Corpus of Contemporary American English (COCA) which contains information from 1990 to 2019 from American English texts. The study was carried out with the purpose of demonstrating key differences in meanings, degrees of formality, and adjective collocations of the two synonymous nouns.

1.6 Significance of the Study

The study of corpus analysis of *scent* and *fragrance* will benefit both teachers and learners in several aspects. First of all, the study will distinguish between the synonyms in terms of degrees of formality and collocations, which may provide guidelines to differentiate them by context of usage. The study also presents current data collecting methods of corpora (COCA) that could be applied to English learning and teaching of synonyms. Lastly, the results of the present study can be applied to language pedagogy in terms of increasing awareness of the differences between words

sharing similar meanings. It cannot be assumed that words in the same set of synonyms can replace each other in every context and the findings of this study may assist teachers and students in correctly using English synonyms.

1.7 Organization of the Study

The paper is organized into five chapters. The first chapter contains the study's background, research purpose, research questions, definitions of key terms, the scope of the investigation, the significance of the study, and study organization. The second chapter provides an overview of the literature on synonyms, criteria for distinguishing synonyms, and previous related studies. The third chapter presents the research instrument, the methodology, and the data analysis. The findings of this study are revealed in chapter four. Finally, chapter five contains the conclusions and suggestions for future research.



CHAPTER 2

REVIEW OF LITERATURE

This chapter reviews the relevant academic literature in five main areas: (1) concept of synonyms, (2) corpus linguistics and language pedagogy, (3) criteria for distinguishing synonyms, (4) English language corpora, and (5) previous related studies.

2.1 Concept of Synonyms

The meaning of *synonym* according to Cruse (2006 as cited in Sirikan and Phoocharoensil, 2014) is words in the same language that mean nearly the same thing as other words.

Oxford Advanced Learners' Dictionary (2010, p. 151) defines a *synonym* as a word or expression that shares with one another the similar or almost the similar meaning in a language. *Small*, for example, is a synonym for *little*.

Kreidler (1998, pp. 97-98) defines *synonyms* as a relationship between two words that share the same features to each other, and the same meanings exist in both words. Moreover, he proposes that they can be distinguished in terms of connotation, dialect and pragmatic value.

Taylor (2003, p. 264) proposes that a *synonym* is an occurrence in which a single meaning is correlated with more than one item. For instance, in English, the adjectives *high* and *tall* exhibit a relationship of co-extension, wherein they convey distinct viewpoints of recessive and dominant. The term *tall* represents the recessive perspective, whereas *high* conveys the dominant perspective. Additionally, it is worth noting that *near synonyms* are more prevalent in English, while *absolute synonyms* are relatively rare.

2.1.1 Absolute Synonyms

Taylor (2003 as cited in Phoocharoensil, 2020, p. 2) states that a pair of synonyms with identical meanings are identified as *absolute synonyms* or *perfect synonyms*, meaning that they can perfectly replace each other in all contexts.

Curse (2006 as cited in Sirikan and Phoocharoensil, 2014, p. 6) also points out that a type of synonyms that barely exists is *absolute synonyms* which are a word having an identical meaning to another word.

Similarly, Fromkin and Rodman (1978, p. 173) argue that an *absolute synonyms* are rarely found. For example, although the words *sofa* and *couch* are considered to have the same meaning, the similarity is actually dependent on a wide range of semantic features.

2.1.2 Near Synonyms

Murphy (2003 as cited in Kiatthanakul, 2015, p.8) defines the meaning of *near synonyms* as words that share same meanings to one another, they only share some central meanings but cannot be used interchangeably for one another in all contexts.

Inkpen and Hirst (2006) give the meaning of *near synonyms* as words that are nearly synonyms but they cannot be totally substitutive for one another. They also mention that it is rare to find *absolute synonyms* available as a dictionary of synonyms or a thesaurus tends to present *near synonyms*.

Hatch and Brown (1995) propose that *near synonyms* do not have the same features as *absolute synonyms*. A study of Phoocharoensil of a pair of English synonyms *mad* and *insane* is a good example. These two words share the core meaning of *mentally ill*. However, the joining of the grammatical pattern of the word *insane* was not likely to happen in native speakers' speech (Phoocharoensil, 2010). Therefore, synonyms cannot be used interchangeably regarding the overlap of their meanings unless the contexts of use are shared (Thornbury, 2002).

2.2 Corpus Linguistics and Language Pedagogy

Hunston (2002, p. 137) states that the applications of the corpus linguistics are beneficial to English language teaching (ELT) in two major areas. Firstly, new descriptions of a language presented by corpora allow language teachers to perceive content that has changed radically. Secondly, the corpora themselves can benefit the production of the language teaching material and syllabus design, out of which the foundation for new language approaches are generated.

Mukherjee (2006) points out that lexicography in language teaching is most significantly impacted by corpus linguistics. A number of English dictionaries, i.e., the Collins Cobuild English Language Dictionary (1987), and Macmillan English Dictionary (2002), describe English words based on the corpus-based descriptions which include information about grammar use for individual words. The following adjective *ashamed* derived from the corpus-based Oxford Advanced Learner's Dictionary (2010) demonstrates lexicographical data that is only available in a corpus:

ashamed /ə'ʃeɪmd/ *adj.* [not before noun] **1** ~ (of sb/sth/ yourself) feeling SHAME or embarrassment about sb/ sth/ or because of sth you have done: *Some patients even feel ashamed of their illness.* **2** ~ **to do sth** unwilling to sth because of SHAME or embarrassment: *She was ashamed to reveal her feelings.*

(OALD, 2010, p. 48)

Regarding the major fields of corpus usage in the context of classroom, data-driven learning (DDL) is mentioned by many corpus linguists as an approach promoting the corpora use in ELT classroom. In fact, the term DDL, introduced by Johns and King (1991), refers to computer-generated concordances used in a classroom setting to expose learners to regular patterns of a language use, as well as task and exercise production for language teaching and learning based on the output of concordances. The use of DDL activities can encourage autonomous learner approaches as the students are allowed to work on corpora themselves such as in corpus-browsing projects and concordance-based exercises that allow learners to explore the use of English synonyms and appropriate noun collocations via observing concordances.

2.3 Criteria for Distinguishing Synonyms

Two criteria, namely collocations and degrees of formality, will be used to distinguish the synonymous nouns, as they can be utilized to investigate the target synonyms *scent* and *fragrance* in terms of similarities and differences.

2.3.1 Collocations

Oxford Advanced Learners' Dictionary (2010) explains that a *collocation* is a word combining with one another in the same language more frequently and regularly

than would happen by chance. Moreover, Smadja (1993 as cited in Chalong and Gampper, 2014, p. 10) defines a *collocation* as a group of words that repeatedly occurs together more frequently than would be expected by coincidence and corresponds to the arbitrariness of word usage.

Interestingly, Seratan (2011 as cited in Sirikan and Phoocharoensil, 2014, p. 7) contents that the definition for *collocation* is still under discussion. He defines *collocation* as typical word combinations that differ from regular combinations in that their components co-occur more frequently than chance would predict in a short stretch of text.

As Conzett et al. (2000) classifies *collocations* into two categories consisting of *strong collocations* and *weak collocations*. The words *winding* and *road* are frequently used together as *winding road*, indicating a *strong collocation*. On the other hand, the weak type has various co-occurrences. The adjective *nice*, for example, can co-occur with various other words such as *nice trip*, *nice day*, *nice smell*, *nice time*, and so on. *Collocations* should be addressed in learning synonyms in order to choose the appropriate words to collocate with each synonym.

2.3.2 Degrees of Formality

Words having similar meanings can be differentiated by degrees of formality. This suggests that certain words are more commonly used in formal contexts compared to others. For instance, according to the Longman Dictionary of Contemporary English (2018), the words *appear* and *reserve* are typically used in formal contexts, whereas *seem* and *keep* are more commonly used in informal contexts.

Jackson and Amvela (2007, as cited in Phoocharoensil, 2021) propose that within groups of synonyms, some words may be used as slang in spoken English, while others are considered standard terms. The following list shows examples of the different degrees of formality between pairs of synonyms.

Formal	Informal
comprehend	get
consume	eat
popular	hot

Formal	Informal
require	need
unwell	sick

(Longman Dictionary of Contemporary English, 2018)

2.4 English Language Corpora

The Corpus of Contemporary American English (COCA) is considered as one of the largest corpora representing American English, and it is probably the most widely used corpus of English. COCA was created by Mark Davies who actually created many other English corpora. These corpora were previously called the BYU Corpora, providing unparalleled insight into English variation.

The Corpus of Contemporary American English (COCA) consists of annually updated data containing one billion words collected between 1990 and 2019. The data primarily originates from five established genres: spoken language, fiction, popular magazines, newspapers, and academic texts. Additionally, three additional genres were added in March 2020, namely TV and movie subtitles, blogs, and other web pages.

The corpus comprises a total of 485,202 texts, with an equal distribution across the eight genres. These texts are sourced from a diverse range of resources, which can be described as follows:

TV/Movies subtitles: The subtitles were taken from the OpenSubtitles collection and are considered more informal than actual spoken data.

Spoken: Over 150 different radio and TV programs, such as All Things Considered (NPR), Hannity and Colmes (Fox), Newshour (PBS), Good Morning America (ABC), etc., were transcribed from unscripted conversations.

Fiction: Children's magazines, first edition books from 1990 until present (first chapters), literary magazines (short stories and plays), popular magazines, and movie scripts.

Popular Magazines: Almost 100 different magazines with a great mix between specific domains, including financial, home and gardening, news, etc. A few examples of the magazines are Cosmopolitan, Fortune, Good Housekeeping, Time, etc.

Newspapers: Ten newspapers across the United State, namely the Atlanta Journal Constitution, USA Today, the New York Times, etc., were consulted with a

good mix between different sections of the newspaper, for example, local news, finance, sports, etc.

Academic Journals: The texts were selected from almost 100 peer-reviewed journals covering the entire range of the Library of Congress classification system.

Blogs: These texts belong to the GloWbE corpus which represents a subset of the texts from the United States. According to the restrictions of blog searchers ruled by Google, almost all the texts were derived from actual blogs.

Web pages: These are representative of a subset of the "General" texts from the United States in the corpus of GloWbE. Some of the texts actually come from blogs because of the search policy from Google that resulted in searches restricted to blogs.

2.5 Previous Related Studies

A number of corpus-based analysis research studies have examined synonyms in both English language and foreign languages. English synonym analysis research will be the main focus of the present study.

Kiatthanakul (2015) investigated the usage differences between three synonyms *achieve*, *accomplish*, and *attain* with regard to meaning and collocation mainly focusing on the usage of the past forms of the synonyms using two dictionaries, namely Oxford Advanced Learner's Dictionary (2015) and Longman Dictionary of Contemporary English (2014). To gather additional insights not found in dictionaries, the Corpus of Contemporary American English (COCA) was consulted by examining the first three hundred concordance lines for each target synonym through random searches. This analysis revealed that the three verbs share common meanings and collocations, indicating that they are near synonyms and can be interchangeably used in certain contexts. Furthermore, the researcher emphasized the significance of introducing corpus-based data to both teachers and students, as it can serve as valuable teaching and learning material for the English language.

In a study conducted by Phitayakorn (2016), an analysis was carried out to examine the patterns of grammar and collocation usage in British English for three verbs: *advise*, *recommend*, and *suggest*. The Longman Dictionary of Contemporary English, 6th edition (2014) and the British National Corpus (BNC) were utilized as data sources to gather information on British English usage. The results revealed that

although the three synonyms share core meanings, they cannot be used interchangeably in all contexts due to differences in grammar patterns and collocations. Consequently, these verbs should be considered near-synonyms. From a language pedagogy perspective, the researcher emphasized the importance of incorporating concordance lines from corpus data in the study of synonyms. It was suggested that L2 learners not overlook the valuable insights provided by corpus data, as dictionaries may not cover all possible patterns found in the corpus. Hence, concordance lines can serve as a supplementary resource, along with dictionaries, for learning synonyms.

Similarly, Gu (2017) conducted a corpus analysis study on the verbs *obtain* and *gain*, focusing on genre, colligation, collocation, and semantic prosody. The online tools Sketch Engine, BNC Web, and Just the Word were used to collect the data. The findings demonstrated that *obtain* is frequently collocated with nouns and tends to appear in passive voice patterns with prepositions, while the verb *gain* is often collocated with abstract nouns and exhibits positive semantic prosody. Furthermore, it was observed that the Oxford Dictionary seemed to lack information on semantic prosody and the frequency of collocational use, such as *obtained by pretense*. The researcher suggested that English foreign language teachers should introduce synonyms to learners by explaining their meanings in traditional practice and include semantic prosody when translating English into Chinese.

In another study, Phoocharoensil (2021) investigated the similarities and differences between the verbal synonyms *persist* and *persevere* in terms of genres, collocations, and semantic preference/prosody. The Corpus of Contemporary American English (COCA) and the Longman Dictionary of Contemporary English, 6th Edition (LDOCE), were employed to collect the data. The results showed that academic texts, followed by webpages, exhibited the highest frequency of occurrence for the two synonymous verbs. Despite their shared meanings, the verbs were found to co-occur with different groups of collocations. In terms of semantic prosody, *persist* frequently co-occurred with negative words or phrases relevant to situations involving successive displeasure, while *persevere* tended to co-occur with lexical items representing the completion of difficult tasks and displaying strong determination with great effort. The corpus evidence also revealed that *persevere* can function as a transitive verb, although learner dictionaries suggest using it only intransitively. Hence, incorporating corpus

data and learner dictionaries in vocabulary instruction allows for greater authenticity in understanding collocational and grammatical patterns.

Lastly, Thongpan (2022) conducted a corpus-based study on synonymous adjectives *far*, *distant*, and *remote*, focusing on their distribution across eight genres and their collocation patterns with semantic preference. The Corpus of Contemporary American English (COCA) was utilized to collect the data. The findings indicated that *remote* had the highest degree of formality among the three adjectives and was predominantly used in academic texts. Regarding collocations and semantic preference, the three synonymous adjectives shared only a few collocates, resulting in varied semantic preferences. It was also observed that *far* was rarely used with adverbs, which differs from its synonyms. Therefore, despite sharing core meanings, these adjectives cannot be used interchangeably in every context. The researcher suggested that the study's outcomes may be useful for English language teachers in designing materials for teaching and learning English synonyms, considering the variation in frequency across genres and the collocations with semantic preference, which can assist learners in distinguishing between synonyms and using them appropriately in various contexts.



CHAPTER 3

RESEARCH METHODOLOGY

This chapter presents: (1) the selected synonyms, (2) the research instruments, and (3) data analysis.

3.1 Selected Synonyms

The target synonyms in this study are two nouns, namely *scent* and *fragrance* as they are mutually defined as *a pleasant smell* or *a pleasant smell that something has* by two online dictionaries: Longman Dictionary of Contemporary English (2018) and Oxford Advanced Learners' Dictionary (2023).

3.2 Research Instruments

This study consulted two learners' dictionaries in the online version, i.e., Longman Dictionary of Contemporary English (LDOCE) and Oxford Advanced Learners' Dictionary (OALD). The selected dictionaries were used to investigate the definitions of the two synonyms.

The Corpus of Contemporary American English (COCA) was also employed in this study as a data source for determining the frequencies of the two synonyms across degrees of formality and adjective collocations. The researcher decided to consult COCA for the genres and adjectives collocates based on four main reasons. First, COCA is considered as one of the largest corpus sources of English language with sufficient patterning of English lexis and grammar which includes more than one billion words of texts updated annual. Second, besides its large size, COCA is available in an online version, which welcomes all users to get access to all data provided without technical conditions of computer use. Third, COCA encompasses eight genres of texts, namely spoken word, fiction, magazines, newspapers, academic texts, TV and movie subtitles, blogs, and webpages. Last, the texts are classified in terms of the length of time from 1990-2020, which makes the analysis more convenient.

3.3 Data Analysis

The target synonyms *scent* and *fragrance* were investigated with regard to their meanings, degrees of formality and adjective collocations. The study began with an exploration of the definitions of the target synonyms by consulting the two online dictionaries. In the following step, the frequencies and distribution through eight genres of the two nouns were derived from COCA, along with an examination of adjective collocates frequently occurring with the two synonymous nouns. In addition to determining the collocates, this study selected the collocates appearing in the top-20 frequency list based on their frequency shown in COCA and whose MI score was over 3, which is the significance value for collocational association (Cheng, 2012 as cited in Phoocharoensil, 2021).



CHAPTER 4

RESULTS

This chapter aims to report the results of the synonymous nouns *scent* and *fragrance* in terms of similarities and differences in the meanings, degrees of formality, and adjective collocations.

4.1 Meanings

Two online dictionaries, namely Longman Dictionary of Contemporary English (LDOCE) and Oxford Advanced Learners' Dictionary (OALD), were used as a main source of the study to investigate this aspect. It revealed that *scent* and *fragrance* shared the same core meaning, i.e., *a pleasant smell that something has or a liquid with a pleasant smell* but differ in detailed meanings. The word *scent* can also refer to *the smell of a particular animal or person that some other animals can follow* which does not exist in *fragrance* as shown in Table 4.1.

Table 4.1

Meanings of 'scent' and 'fragrance' from LDOCE (2018) and OALD (2023) in the online versions

Synonyms	Meanings	
	LDOCE	OALD
scent	1. [countable] a pleasant smell that something has e.g., a yellow rose with a lovely <i>scent</i> scent of e.g., the sweet <i>scent</i> of ripe fruit 2. [countable] the smell of a particular animal or person that some other animals, for example dogs, can follow e.g., The deer caught the <i>scent</i> of the man.	1. [uncountable, countable] the pleasant smell that something has e.g., The flower has no <i>scent</i> . 2. [uncountable, countable, usually singular] the smell that a person or an animal leaves behind and that other animals such as dogs can follow e.g., The hounds have the <i>scent</i> . 3. [uncountable] (especially British English) a liquid with a pleasant smell that you wear on your skin to make it smell nice

Synonyms	Meanings	
	LDOCE	OALD
	3. [countable, uncountable] especially British English a liquid that you put on your skin to make it smell pleasant e.g., Was the <i>scent</i> essentially McAllister, or was it some subtle perfume she chose to wear?	e.g., She dabbed some <i>scent</i> on her neck.
fragrance	1. [countable, uncountable] a pleasant smell fragrance of e.g., the rich <i>fragrance</i> of a garden flower 2. [countable] a liquid that you put on your body to make it smell pleasant e.g., Would you like to try White Diamonds, the new <i>fragrance</i> from Elizabeth Taylor?	1. [countable, uncountable] a pleasant smell e.g., The bath oil comes in various <i>fragrances</i> . 2. [countable] a liquid that you put on your skin in order to make yourself smell nice e.g., an exciting new <i>fragrance</i> from Dior

4.2 Degrees of Formality

The Corpus of Contemporary American English (COCA) was consulted to explore the degrees of formality across eight genres, as demonstrated in Table 4.2.

Table 4.2

Total frequency and division of 'scent' and 'fragrance' through eight genres

Genre	Scent		Genre	Fragrance	
	Frequency	Per million		Frequency	Per million
fiction	4,919	41.57	magazines	1,209	9.59
magazines	2,169	17.20	fiction	707	5.98
TV and movie subtitles	912	7.12	newspapers	346	2.84
webpage	544	4.38	blogs	327	2.54
newspapers	531	4.36	webpage	176	1.42
blogs	523	4.07	TV and movie subtitles	156	1.22
spoken	289	2.29	spoken	134	1.06
academic texts	252	2.10	academic texts	84	0.07
Total	10,139		Total	3,139	

The total frequency of *scent* (10,139 tokens) is over three times higher than that of *fragrance* (3,139 tokens) and occurs with a higher frequency than *fragrance* in each genre presented by COCA. Regarding the degrees of formality, the two synonyms tend to appear more often in informal contexts, which can be seen in the top four genres, while academic texts — which represent formal styles or contexts — had the lowest frequency for both synonyms. Moreover, if we consider the highest frequency of occurrences of both synonyms, the two nouns are commonly found in written English, as they are mostly used in fictions and magazines.

4.3 Adjective Collocations

In order to investigate the adjective collocations of the target synonyms, this study obtained the data from the Corpus of Contemporary American English (COCA). The collocates of *scent* and *fragrance* were selected based on their frequency in the top-20 list with a MI score of ≥ 3 , as shown in Table 4.3.

Table 4.3

Adjective collocates of 'scent' and 'fragrance' from COCA

Rank	Scent			Fragrance		
	Adjective collocate	Frequency	MI-Scores	Adjective collocate	Frequency	MI-Scores
1	sweet	321	5.66	sweet	115	5.74
2	fresh	185	4.68	favorite	34	4.07
3	faint	168	7.26	floral	34	8.20
4	strong	158	3.13	fresh	31	3.66
5	familiar	101	4.31	spicy	30	7.54
6	rich	97	3.54	heavy	27	3.44
7	heavy	97	3.73	heady	27	8.68
8	warm	94	4.07	light	22	3.47
9	floral	83	7.93	wonderful	21	3.20
10	clean	80	3.84	pleasant	21	5.40
11	heady	79	8.68	artificial	17	5.15
12	pungent	75	9.07	warm	16	3.07
13	spicy	67	7.15	delicious	16	5.02
14	musky	63	10.85	delicate	16	5.25
15	sharp	53	4.05	synthetic	15	6.15
16	pleasant	52	5.16	fruity	14	8.56
17	lingering	52	6.71	delightful	12	6.25
18	acidic	48	9.51	scented	12	8.50

Rank	Scent			Fragrance		
	Adjective collocate	Frequency	MI-Scores	Adjective collocate	Frequency	MI-Scores
19	thick	45	3.54	lovely	11	3.55
20	burning	45	5.01	sexy	11	4.20

The findings show that the target synonymous nouns share eight adjective collocates, including *floral*, *fresh*, *heady*, *heavy*, *pleasant*, *spicy*, *sweet*, and *warm*. Among the adjective collocates, *sweet* had the highest frequency for both synonymous nouns. However, the adjectives *rich* and *strong* also co-occur with the two synonyms in COCA. The two adjectives were excluded from the *fragrance*'s list due to their low MI scores of 2.24 and 2.27, respectively, which did not reach the target score of ≥ 3 . Moreover, the adjective *sensitive* is also in the top-20 frequent collocates of *fragrance*, with a MI score of 3.63. The word was omitted from the list according to its irrelevant collaboration with *fragrance*. Rather, the adjective tends to serve the function of a subject complement, as shown in the concordance lines by COCA below:

- 1) her co-workers were **sensitive** to **fragrance** and prone to asthma attacks ...
- 2) ... you're **sensitive** to **fragrance**, she recommends diluting essential oils, ...
- 3) But I'm **sensitive** to **fragrances**, which presents a problem.
- 4) ... skin is thicker and generally less **sensitive** to irritation by **fragrances**, ...
- 5) ... a new aftershave, and Japanese people are **sensitive** to foreign **fragrances**.

CHAPTER 5

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

This chapter presents (1) summary of the findings, (2) discussion of the two synonyms *scent* and *fragrance* in terms of similarities and differences, (3) the conclusion, (4) the implications, and (5) the recommendations for future research.

5.1 Summary of the Findings

The purpose of this study was to examine the two synonymous nouns *scent* and *fragrance* in terms of similarities and differences in meanings, degrees of formality, and adjective collocations by consulting two online dictionaries, namely Longman Dictionary of Contemporary English (LDCE) and Oxford Advanced Learners' Dictionary (OALD), and a massive corpus, namely the Corpus of Contemporary American English (COCA). The results are summarized as follows:

Firstly, the two synonyms share the same core meaning of *a pleasant smell something has* and *a liquid that is used to put on body or skin to make it smell pleasant* so that *scent* and *fragrance* can be classified as near-synonyms. Regarding the differences of the synonyms based on the selected dictionaries, the second meaning of *scent* which is relevant to *the smell of a person and animal*, is not applied to *fragrance*. As a result, there is no chance that the synonyms can be substituted for each other in all contexts since the synonyms do not share all their features.

Secondly, the target synonyms tend to appear in less formal contexts more often than formal contexts, as indicated by their top four genres, namely fiction, magazines, TV and movie subtitles, and web page from *scent*'s list; and magazines, fiction, newspapers, and blogs from *fragrance*'s list, respectively. According to the highest frequency of occurrences of both synonyms, *scent* and *fragrance* are the most used in written English, with fiction and magazines representing informal written texts. On the other hand, both synonyms are least frequently used in spoken and academic texts.

Thirdly, the two nouns *scent* and *fragrance* share eight adjective collocates from their top-20 frequency list: *floral*, *fresh*, *heady*, *heavy*, *pleasant*, *spicy*, *sweet*, and

warm. The collocations belonging to both nouns all had a MI score of 3 and above; otherwise, they would have been excluded from the lists.

Finally, the selected online dictionaries, namely Longman Dictionary of Contemporary English (LDCE) and Oxford Advanced Learners' Dictionary (OALD), provide only information about the meanings of the two synonyms, which is not enough to observe the similarities and differences. As a consequence, the English corpus, namely the Corpus of Contemporary American English (COCA), was used to investigate degrees of formality and adjective collocations in order to distinguish the synonyms.

5.2 Discussion

The nouns *scent* and *fragrance*, which are considered near-synonyms, share core meanings. However, dictionaries indicate that *scent* can also refer to *the smell of an animal or person*, which may not necessarily be pleasant, whereas this specific meaning does not exist in the definitions of *fragrance*. These findings support Hatch and Brown's theory (1995) that near-synonyms cannot share all their features like absolute synonyms. Therefore, the words *scent* and *fragrance* cannot be used interchangeably in all contexts, as Thornbury (2002) confirms that although there is overlap in their meanings, they cannot replace each other unless the specific contexts of use are shared.

Analyzing the total frequency of occurrences of these two nouns across eight genres in COCA, *scent* appears 10,139 times, while *fragrance* appears 3,139 times, which is three times lower than the frequency of *scent*. This implies that the word *scent* is used more frequently than the word *fragrance* in the American corpus. Furthermore, it can be assumed that due to its wider range of meanings, *scent* tends to appear in a broader context where *fragrance* cannot be applied. The following concordance lines from COCA exemplify contexts where only *scent* can be used because of its definition, which refers to *the smell of a specific animal or person* that cannot be replaced by *fragrance*.

- 1) ... approached because the snake would catch my **scent** and turn away.
- 2) The good **scent** of hay and horse rushed to his face as he stepped inside ...
- 3) ... **scent** of their mother's breast milk was enough to calm the newborns ...

- 4) ... for **scent** trails that lead to sea lions or their relatives then wait along ...
- 5) Then he sniffed the air like an old dog on a **scent**.

Regarding the top 20 adjective collocations of the target synonyms, there are several instances where the target synonym *scent*, particularly, co-occurs with negative adjectives. These negative adjectives include *burning*, *acidic*, *heady*, *heavy*, *pungent*, *sharp*, *strong*, and *thick*. In such cases, the original core meaning of *a pleasant smell* is transformed into *an unpleasant or undesirable smell*. The concordance lines from COCA below illustrate this transformation:

- 1) I wrinkled my nose at the **pungent scent** of vinegar, which my mother ...
- 2) ... as I stood there in the afternoon heat with the **sharp scent** of juniper ...
- 3) ... the sharp **acidic scent** that make Jim start coughing uncontrollably.
- 4) You don't want that **strong fragrance**. That's very old-fashioned.
- 5) I wouldn't want to disturb a baby's sense of smell with **heavy fragrances**.

Based on the research findings, it appears that nouns with inherently positive meanings can be used to express negative connotations when they occur with negative collocations. An example of this phenomenon is the English noun *success*, which undergoes a shift in meaning from *achieving positive outcomes* to *not achieving the desired results* when it co-occurs with negative adjectives like *limited*, *modest*, and *partial*. This change in meaning is demonstrated by the following concordance lines from COCA:

- 1) ..., but with **limited success**, because complicated curves are hard to ...
- 2) it was first released to the public in 1972 and It only met with **limited success**.
- 3) **Modest success** was followed by crushing failure, creating an uncertainty ...
- 4) ... after World War II was only a **partial success**. The Libyan monarchy ...
- 5) We have to look for this **partial success** before we achieve the complete ...

5.3 Conclusion

The synonyms *scent* and *fragrance* can be considered near-synonyms as they share core meanings, but they are not interchangeable in every context due to the additional definition of *scent* that refers to *the smell of animals or humans*, which is not present in *fragrance*. When it comes to formality, both synonyms tend to be used more frequently in informal contexts, as evidenced by their prevalence in genres such as

fiction and magazines, which are typically associated with written English. Analyzing the top 20 frequent collocates of these synonyms, we observe that eight adjective collocates are shared between them, including *floral*, *fresh*, *heady*, *heavy*, *pleasant*, *spicy*, *sweet*, and *warm*. It is worth noting that there is an interesting occurrence of negative adjectives associated with the target synonyms, particularly *scent*, indicating *an unpleasant or unwanted smell*, such as *acrid* or *pungent scent*.

5.4 Implications

The present study has demonstrated the value of utilizing both dictionaries and corpora in English vocabulary learning and teaching for second language learners and English language teachers. Dictionaries serve as a primary tool for accessing brief information about vocabulary, while corpora offer more in-depth details that facilitate word comparisons. Therefore, consulting the corpus is essential for exploring aspects of vocabulary that dictionaries may not cover comprehensively, such as word frequency across different genres and collocations with their corresponding frequencies, providing a reflection of authentic language usage by English language users.

The findings highlight that words with similar meanings cannot always be used interchangeably in all contexts. Consequently, language learners and teachers need to be aware of the nuanced differences between words that share similar meanings. For language teachers, incorporating corpus demonstrations in the classroom becomes crucial to encourage learners to delve into more comprehensive information about English words that might not be readily available in dictionaries. The integration of corpus-based activities and learning materials can facilitate this process. For instance, learners can be encouraged to explore the appropriate adjective collocations for synonyms and engage in corpus-browsing projects to observe concordances.

Language learners, on the other hand, must recognize that relying solely on dictionaries for understanding synonyms is insufficient. Dictionaries often provide limited examples and might exclude in-depth information to avoid overwhelming users with extensive data. Therefore, learners should incorporate the use of corpora in their vocabulary learning endeavors, as corpora offer up-to-date data and depict real-world English language use.

5.5 Recommendations for Future Research

Based on the findings and conclusions of this study, two recommendations are proposed for future research. Firstly, conducting an analysis of the grammatical patterns and semantic prosody of adjective collocations for the synonyms *scent* and *fragrance* would provide more comprehensive insights into the possible patterns of grammar and contextual usage in which these synonyms co-occur with their collocates. Examining grammatical patterns and semantic prosody can be considered additional criteria for distinguishing synonyms, thereby enhancing the validity of the findings through the application of multiple criteria in data analysis. Secondly, it is recommended to compare data obtained from two English corpora that represent both American English and British English. This comparative analysis would explore the similarities and differences between *scent* and *fragrance* in these two major varieties of standard English. Such a comparison has the potential to offer new insights into the criteria used to distinguish English synonyms and shed light on potential variations between the two linguistic contexts. These future research directions hold promise for enriching our understanding of *scent* and *fragrance* as synonyms, and their distinct usage patterns across different grammatical and semantic contexts, as well as within different English language varieties.



REFERENCES

- Brezina, V., & Gablasova, D. (2018). The corpus method. In J. Culpeper, P. Kerswill, R. Wodak, T. McEnery, & F. Katamba (Eds.), *English language: Description, variation and context* (2nd ed.). Palgrave Macmillan.
- Cheng, W. (2012). *Exploring corpus linguistics: Language in action*. Routledge.
- Conzett, J., Hargreaves, P. H., Hill, J., Lewis, M., & Woolard, G. C. (2000). *Teaching collocation: Further developments in the lexical approach* (Vol. 244). (M. Lewis, Ed.). Language Teaching Publications.
- Cruse, A. (2006). *Glossary of semantics and pragmatics*. Edinburgh University Press.
- Fromkin, V., & Rodman, R. (1978). *An introduction to language* (2nd ed.). Holt.
- Granger, S. (2004). Computer learner corpus research: Current status and future prospects. *Applied Corpus Linguistics*, 123-145.
- Gu, B. J. (2017). Corpus-based study of two synonyms: Obtain and gain. *Sino-US English Teaching*, 8, 511-522.
- Hunston, S. (2002). Pattern grammar, language teaching, and linguistic variation. *Using Corpora to Explore Linguistic Variation*, 167-183.
- Inkpen, D., & Hirst, G. (2006). Building and using a lexical knowledge base of near-synonym differences. *Computational Linguistics*, 32(2), 223-262.
- Jackson, H., & Amvela, E. Z. (2000). *Words, meaning and vocabulary: Meaning relations*. Atheneum Press.
- Jackson, H., & Amvela, E. Z. (2007). *Words, meaning and vocabulary: An introduction to modern English lexicology*. Bloomsbury.
- Jarunwaraphan, B., & Mallikamas, P. (2020). A corpus-based study of English synonyms: Chance and opportunity. *REFlections*, 27(2), 218-245.

- Jirananthiporn, S. (2018). Is this problem giving you trouble? A corpus-based examination of the differences between the nouns problem and trouble. *Thoughts*, 2018 (2), 1-25.
- Kayaoglu, M. (2013). The use of corpus for close synonyms. *Journal of Language and Linguistic Studies*, 9(1), 128-144.
- Kiatthanakul, M. T. (2015). *A corpus-based study of English synonyms: Achieve, accomplish, and attain* [Master's independent study, Thammasat University].
- Kreidler, C. W. (1998). *Introducing English semantics*. Psychology Press.
- Liu, D., & Espino, M. (2012). Actually, genuinely, really, and truly: A corpus-based behavioral profile study of near-synonymous adverbs. *International Journal of Corpus Linguistics*, 17(2), 198-228.
- Mayor, M. (Ed.). (2009). *Longman dictionary of contemporary English*. Pearson Education.
- Mukherjee, J. (2006). Corpus linguistics and language pedagogy: The state of the art and beyond. *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*, 5-24.
- Murphy, M. L. (2003). *Semantic relations and the lexicon: Antonymy, synonymy and other paradigms*. Cambridge University Press.
- O'keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.
- Oxford. (2010). *Oxford advanced learners dictionary (dictionaries intermediate to advanced)*. Oxford University Press.
- Oxford. (2014). *Oxford advanced learners' dictionary of academic English*. Oxford University Press.
- Oxford. (2015). *Oxford advanced learners' dictionary*. Oxford University Press.

- Partington, A. (1998). *Patterns and Meanings*. John Benjamins B.V.
- Petcharat, N. & Phoocharoensil, S. (2017). A corpus-based study of English synonyms: Appropriate, proper, and suitable. *LEARN Journal: Language Education and Acquisition Research Network*, 10(2), 10-24.
- Phitayakorn, M. W. (2016). *A corpus-based study of English synonyms: Advise, recommend, and suggest* [Master's independent study, Thammasat University].
- Phoocharoensil, S. (2010). A corpus-based study of English synonyms. *International Journal of Arts and Sciences*, 3(10), 227-245.
- Phoocharoensil, S. (2020). A genre and collocational analysis of consequence, result, and outcome. *3L: The Southeast Asian Journal of English Language Studies*. 263(3), 1-16.
- 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.
- Ruenroeng, C. (2014). *A corpus-based analysis of English synonyms: Ruin, demolish, destroy*. [Master's independent study, Thammasat University].
- Seretan, V. (2011). *Syntax-based collocation extraction*. Springer.
- Sinclair, J. (1999). *Concordance tasks*. Online at <http://www.twc.it/happen.html> (last access: February 2023).
- Sirikan, N. (2014). *A corpus-based study of English synonym: Appalling, horrific, horrible, hideous*. [Master's independent study, Thammasat University].
- Smadja, F. (1993). Retrieving collocations from text: Xtract. *Computational Linguistics*, 19(1), 143-178.

- Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language*, 2(1), 23-55.
- Taylor, J. R. (2003). Near synonyms as co-extensive categories: 'High' and 'Tall' revisited. *Language Sciences*, 25(3), 263-284.
- Thongpan, N. (2022). A corpus-based study of English synonyms of the adjectives 'Far', 'Distant', And 'Remote'. *Journal of Positive School Psychology*, 3986-4001.
- Thornbury, S. (2002). *How to teach vocabulary*. Longman.
- Tribble, C. (2000). Practical uses for language corpora in ELT. *A Special Interest in Computers: Learning and Teaching with Information and Communications Technologies*, Whitstable, Kent, UK: IATEFL, 31-41.
- Uba, S. Y. (2015). A corpus-based behavioural profile study of near-synonyms: important, essential, vital, necessary and crucial. *International Journal of English Language and Linguistics Research*, 3(5), 9-17.