



Is it ‘important,’ ‘significant,’ or ‘crucial’? A Corpus Based Study of English Synonyms

Wimonnit Chaokongjakra

wimonnit.c@litu.tu.ac.th, Language Institute, Thammasat University, Thailand

APA Citation:

Chaokongjakra, W. (2023). Is it ‘important,’ ‘significant,’ or ‘crucial’? A corpus based study of English synonyms. *LEARN Journal: Language Education and Acquisition Research Network*, 16(2), 512-532.

Received
20/05/2023

Received in revised
form 13/06/2023

Accepted
20/06/2023

ABSTRACT

Due to the large number of near-synonyms present in the English language, English learners frequently struggle to use near-synonyms in different contexts, as these words, despite similar meanings, are not always interchangeable. This study examines the distribution and collocation of three synonyms, ‘important,’ ‘significant,’ and ‘crucial,’ across genres. The objectives are to investigate differences in formality based on genre distribution and to analyze collocates in relation to semantic preference. The data from the Corpus of Contemporary American English (COCA) indicates that the synonyms are used primarily in academic contexts, with ‘important’ being the most common. Nonetheless, these synonyms differ in their collocates and themes. ‘Significant’ is frequently associated with quantity-related matters, whereas ‘crucial’ is typically associated with the political domain. ‘Important’ often appears with intensity-related adverbs, and ‘significant’ with study-related ones. Therefore, the analysis of collocates between nouns and adverbs reveals that these synonymous adjectives have both shared and distinct preferences.

Keywords: near-synonyms, collocations, genre, semantic preference, COCA

Introduction

English is a necessary language for international communication in areas such as education, business, and entertainment as a global lingua franca. According to Crystal (2007), English has a sizable vocabulary because of its long history and globalization. Wilkins' (1972) famous quote, "Without grammar, very little can be conveyed; without vocabulary, nothing can be conveyed," proves the significance of vocabulary understanding and knowledge. With regard to vocabulary learning, Carter and McCarthy (1988) argued that regardless of how proficient students are in grammar or how well they master the sounds of a second language (L2), communication in an L2 cannot be meaningful without words that express a variety of meanings. Likewise, Folse (2004) noted that learners who do not understand vocabulary are unlikely to comprehend or acquire it. Although vocabulary, according to Gass and Selinker (2008), is critical to language learning, it is also viewed as "the fuel of language" by Gardner (2013). Knowledge of it is essential for learners' development of the four main language skills (Nation, 2013). A growing body of literature recognizes the importance of vocabulary in language learning, especially the use of synonyms to expand vocabulary. However, as Laufer (1997) pointed out, learning English language vocabulary can be challenging because it has an infinite number of synonyms, which makes it both challenging and confusing.

As stated by Carter (2012), one of the most challenging aspects of English language acquisition for L2 learners is the lexical relationship between two or more linguistic forms with the same meaning, known as synonymy. While some synonyms share semantic similarities, using the wrong synonym can affect the naturalness of L2 English production (e.g., '*consequence*' and '*result*' are synonyms, but '*consequence*' is related to words of a negative sense, whereas '*result*' has no negative connotations), as noted by Phoocharoensil (2020a). This indicates that not all words in a set of synonyms are interchangeable within varying contexts.

It is undeniable that the majority of language learners struggle to differentiate the various meanings and applications of synonyms, leading to confusion (Aroonmanakun, 2015). According to Gass and Selinker (2008), English learners who are proficient in a wide range of English vocabulary typically have a significant advantage when producing the language. To master English vocabulary, learners must learn how to use synonyms correctly, as varied word choices can help learners avoid repetition in their writing. The ability to use varied English vocabulary not only helps in conducting daily conversation but also improves knowledge of lexical features and sentence construction, providing more choices for appropriate vocabulary in each context. The definition of synonym as provided by Merriam-Webster Dictionary (online version) is, "one of two or more words or expressions of the same language that have the same or nearly the same meaning in some or all senses." According to Thornbury (2002), no near-synonyms are identical in every respect, and replacing one with its synonym may result in grammatical errors in L2 learners.

This study investigates the synonyms of the adjectives '*important*,' '*significant*,' and '*crucial*' across genres and identifies noun and adverb collocations in relation to the semantic preference of these three target synonyms. Stubbs (2001) defines semantic preference as "the relation, not between individual words, but between a lemma or word form and a set of semantically related words". As an example, the noun collocates that follow the verb '*undergo*' point out a diverse range of semantic preferences, such as medical terms (treatment, hysterectomy, operation), testing (examination, training), and change (changes, transformations). Similarly, semantic preference is the restriction of lexical item co-occurrences to those that share a semantic feature (Sinclair, 2004). In other words, the determination of semantic preference is possible through analyzing the semantic connections between words and their collocates. By using a list of collocations, one can identify the various associations of the target word as well as the semantic relationships between its collocates (Ly & Jung, 2015). Studies such as those by Harley (2006), Chung (2011), and Phoocharoensil (2020a), have shown that despite the fact that synonyms have similar meanings,

they cannot completely replace one another because of differences in context and/or perspectives. The meanings of the three target adjective synonyms ‘*important*,’ ‘*significant*,’ and ‘*crucial*,’ were sufficiently similar to cause confusion among EFL learners and teachers.

In addition, because of their frequency and importance for learners, these adjectives were chosen as the target synonyms. The adjectives ‘*important*,’ ‘*significant*,’ and ‘*crucial*’ are included on Longman Communication 3000 list, with ‘*important*’ ranking among the top 1,000 most frequent words in written and spoken English, ‘*significant*’ among the top 1,000 most frequent words in written English and the top 2,000 most frequent words in spoken English, and ‘*crucial*’ among the top 2,000 most frequently used words in written English. All of these target synonyms are also included in the Oxford 3000 (American English), a list of the 3,000 most important words to learn in English that serves to develop a core word list for advanced English learners.

Table 1 displays the definitions of the three target synonyms according to three online American-English dictionaries: Longman Advanced American Dictionary (online version), Oxford Advanced American Dictionary (online version), and Merriam-Webster Dictionary (online version).

Table 1

Definitions and examples of ‘*important*,’ ‘*significant*,’ and ‘*crucial*’ from American-English dictionaries

	Longman Advanced American Dictionary	Oxford Advanced American Dictionary	Merriam-Webster Dictionary
important	an important event, decision, problem etc. has a big effect or influence on people’s lives or on events in the future <i>e.g., I have an important announcement to make, so please listen carefully.</i>	having a great effect on people or things; of great value <i>e.g., Money played an important role in his life.</i>	marked by or indicative of significant worth or consequence: valuable in content or relationship <i>e.g., In his editorial, he made several important points.</i>
significant	having an important effect or influence, especially on what will happen in the future <i>e.g., There has been a significant change in the tone of the media’s coverage.</i>	large or important enough to have an effect or to be noticed <i>e.g., There are no significant differences between the two groups of students.</i>	having or likely to have influence or effect: IMPORTANT <i>e.g., The study found a statistically significant decrease in symptoms in patients who had taken the drug.</i>
crucial	something that is crucial is extremely important, because everything else depends on it <i>e.g., Ritchey worked for some very good bosses, but it was the difficult boss who had a crucial impact on his career.</i>	extremely important, because it will affect other things <i>e.g., Parents play a crucial role in preparing their children for school.</i>	important or essential as resolving a crisis: IMPORTANT, SIGNIFICANT <i>e.g., Vitamins are crucial for maintaining good health.</i>

From the definitions presented above of ‘*important*,’ ‘*significant*,’ and ‘*crucial*,’ it is clear that these three adjectives can be considered near-synonyms as they have similar meanings of conveying the idea of something being necessary, noteworthy, or something being of great importance or value and cannot be ignored or underestimated. However, these definitions offer only a basic understanding of the meaning and examples of the three near-synonyms. Learners require a more thorough understanding of their usage to use them appropriately in different contexts. Dictionaries, including online dictionaries, do not provide enough information on their

usage in specific contexts, which can lead to incorrect usage by learners (Ly & Jung, 2015). Furthermore, explanations of their formality and collocations are not always clear. Therefore, it is important for L2 learners to understand the similarities and differences of these high-frequency synonymous adjectives. Unfortunately, there is limited research on these three synonyms. To fill this gap, this study aims to analyze the formality and collocational patterns of *'important,' 'significant,'* and *'crucial'* through a corpus-based study (Phoocharoensil, 2020a). By using language corpora, which are collections of normal texts that show how language is used, this study aims to provide insights that are not available from dictionaries. This is due to the fact that limitations of space and conciseness of explanation make it unlikely that a dictionary will include all linguistic information regarding a word. Thus, this study aims to address this issue and provide a more comprehensive understanding. The following section will review the relevant literature on the concept of synonymy, including the two major types of synonyms, methods for distinguishing synonyms based on various criteria, and previous corpus-based studies on synonyms.

Literature Review

Synonymy

Synonymy is a significant yet complex concept that refers to the relationship between various lexical terms. It is a fundamental principle in both lexicology and language instruction. Despite the fact that synonyms are frequently identified through substitution, neither substitution nor dictionary definitions can provide quantitative meaning in terms of data such as sentence structure, collocation, and frequency of structure patterns. There are no perfect synonyms, which makes selecting words difficult. As defined by Taylor (2002), synonymy is 'the phenomenon whereby a single meaning is associated with more than one distinct lexical item'. In his study of the two near-synonymous adjectives *'high'* and *'tall'* in terms of their varying interpretations of an entity's verticality and the co-extension relation, *'high'* designates the dominant vantage, whereas *'tall'* designates the recessive vantage. That is, we can use *'tall'* girl rather than *'high'* girl, despite the fact that these terms are somewhat synonymous. This leads to the two most common characteristics of synonyms: strict or absolute synonyms and loose or near-synonyms based on the degree to which the words overlap in a particular context.

Strict or absolute synonyms refer to words with identical meanings that can be used interchangeably in all sentence contexts without a change in meaning, style, or connotation (Dvorak & Dawson, 2011; Edmonds & Hirst, 2002). In other words, interchangeability between two words indicates that one word can be substituted for the other without affecting the message's meaning or tone. However, such synonyms are quite rare as they are redundant in language (Phoocharoensil, 2010). As Taylor (2002, p. 246) concludes, perfect synonymy is vanishingly rare, methodologically proscribed, or a logical impossibility, what we frequently do encounter are pairs of words that are "near" synonyms.'

By contrast, loose or near-synonyms are words that have similar meanings but are not identical and cannot be used interchangeably in all contexts. Although the meanings of such words frequently overlap, they cannot always be used interchangeably due to factors such as differences in connotation, usage patterns, and variations in meaning and context. Phoocharoensil (2020b) provided examples of near-synonyms, such as *'repair'* and *'mend'*. While these words can be used interchangeably in some contexts, the word *'mend'* sounds more appropriate in the context of clothes, as shown in (2) below.

- (1) He will ***repair/mend*** the bike.
- (2) She will carefully ***mend/*repair*** the tear in her shirt.

To summarize, loose or near-synonyms are more common than strict or absolute synonyms, and using them incorrectly can affect the meaning, style, and tone of a text. It is

important for ELT teachers and learners to be able to differentiate between near-synonyms based on a variety of criteria, such as formality or the style of the context, connotation, collocational patterns, and semantic preference (Phoocharoensil, 2010). As this study investigates two major issues, formality and collocation, they will each be discussed in the following section.

Distinguishing synonyms

Various criteria are employed to distinguish synonyms, such as formality, connotation, collocational pattern, and semantic preference (Phoocharoensil, 2010; Phoocharoensil, 2020a). These criteria play a part in distinguishing between near-synonyms so that the appropriate vocabulary can be used in a given context. Although the two words have a similar meaning, one may be more appropriate for formal settings while the other is more appropriate for informal settings. Formal words are typically found in formal contexts, such as academic texts, whereas informal words are typically found in informal contexts, such as spoken language, everyday conversations, and novels. However, as Nisani (2015) added, most dictionaries do not indicate a word's formality, which depends on its context. To understand the degree of formality between synonyms, background knowledge and practice with the target language are required. As noted in Petcharat & Phoocharoensil (2017)'s study, '*appropriate*' is generally considered the most formal, while its synonym '*proper*' is the least formal. '*Suitable*' is situated between the two. Additional examples of English words that differ in their degree of formality are provided below.

Formal Language

commence
terminate
endeavor

Informal Language

start
end
try

(Cambridge Dictionary, online version)

Another crucial criterion for distinguishing near-synonyms is collocation. John Rupert Firth, the father of collocation, introduced it in 1957. According to Firth, a collocation is a combination of frequently occurring words (Martynska, 2004). This means that some synonyms are expected to appear together with certain words. Typically, collocations are evaluated through statistical techniques such as mutual information (MI) scores, Z-scores, log-likelihood, and other similar methods (Saito, 2020). According to Longman Dictionary of Contemporary English (online version), collocation is defined as 'the way in which some words are often used together, or a particular combination of words used in this way'. It provides the example of 'commit a crime' as a typical collocation in English. Phoocharoensil (2022) provides further clarification with the words '*rancid*' and '*addled*'. While both '*rancid*' and '*addled*' can refer to spoiled food, their collocates—the words that typically come after them—distinguish these two synonyms. In English, it is more common to use the collocations '*rancid bacon*' and '*addled eggs*' rather than '**rancid eggs*' and '**addled bacon*'. To use near-synonyms correctly in various contexts, it is necessary to consider the criteria for distinguishing them. It is important to note that synonyms can differ when paired with particular collocates. To illustrate, previous corpus-based studies related to synonyms are reviewed in the next sections.

Previous Studies on Synonyms

Numerous studies have been conducted on synonyms, focusing on their similarities and differences. These studies have employed corpus-based data from various corpora, including the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). The most prevalent criteria for distinguishing synonyms are their meanings and senses of meaning, collocations, grammatical patterns, and degrees of formality. These corpus-informed data provide evidence of a word's connotations, degree of formality, and collocational patterns, which is useful

for both language learners and English language teachers in selecting the appropriate word for a given context. According to Chung (2011), corpus-based lexical studies were conducted in this field because they are considerably more reliable than traditional descriptive research.

The study by Petcharat and Phoocharoensil (2017) is an example of a corpus-based investigation into the definitions and applications of three synonymous English adjectives. The researchers focused specifically on the adjectives *'appropriate,' 'proper,'* and *'suitable,'* which are frequently used interchangeably in everyday speech. The purpose of the study was to compare and contrast these terms in terms of their meanings, collocations, levels of formality, and grammatical structures. The researchers gathered data from three dictionaries (Longman Dictionary of Contemporary English, Longman Advanced American Dictionary, and Macmillan Collocations Dictionary) and the Corpus of Contemporary American English (COCA) to accomplish this. They then analyzed the data to identify usage patterns and variations. One of the most important findings of the study was that, despite having similar core meanings, the three adjectives are used differently in different contexts. For instance, *'appropriate'* appears to be the most formal term, while *'proper'* appears to be the least formal. However, *'suitable'* seems to fall between *'appropriate'* and *'proper'*. In addition, the three terms have distinct collocates, or words that appear frequently with them. *'Appropriate'* is frequently associated with terms such as *'behavior,' 'manner,'* and *'response,'* whereas *'proper'* is more frequently associated with terms such as *'procedures,' 'nutrition,'* and *'name.'* This is most likely because *'proper'* covers more meanings than *'appropriate'* and *'suitable'*. The study demonstrates the value of using corpus analysis to investigate the variations of language usage and help learners gain a deeper understanding of how words are employed in context.

In Aroonmanakun (2015)'s study, COCA corpus was used to compare and contrast the English synonyms *'quick'* and *'fast'*. The study revealed that while the basic definitions of the two terms overlap, their specific meanings differ in certain contexts. Furthermore, it was shown that the two terms are not always interchangeable and that their usage is context-dependent. While *'quick'* tends to emphasize the action's quality, *'fast'* emphasizes its manner or speed. For instance, a *'quick learner'* is someone who has the ability to learn quickly, whereas a *'fast learner'* is someone who acquires knowledge rapidly. Similarly, *'quick access'* and *'fast access'* may both refer to an action that is completed in a very short amount of time, but *'quick access'* tends to emphasize the short amount of time spent on the activity, while *'fast access'* emphasizes the great speed with which the activity or process is completed. The study also found that the ratio of usage of *'quick'* and *'fast'* with the same nouns shows that *'quick'* is more commonly used overall. *'Quick'* is frequently used with nouns derived from verbs to indicate a short time period, such as answer, calculate, and call, whereas *'fast'* is used to indicate the high speed of the action being performed by that noun, such as acceleration, attack, and connection. However, the study also revealed that the two terms can be used interchangeably in certain contexts, such as in the phrases *'make a quick buck'* and *'make a fast buck,'* both of which convey the same meaning of *'earn money quickly, often dishonestly.'* Overall, the study emphasizes the significance of context and variation when employing synonyms in language learning and communication.

Bergdahl (2009) analyzed three English adjectives, *'beautiful,' 'handsome,'* and *'good-looking,'* utilizing data from COCA and five contemporary dictionaries. Although all three adjectives convey the same basic meaning of being visually appealing, their usage patterns differ. *'Beautiful'* is typically used to describe women, whereas *'handsome'* is typically used to describe men. The term *'attractive'* is gender-neutral according to dictionaries, but in COCA samples it is more frequently used to describe men. Additionally, the term *'handsome'* is sometimes used to describe women who have a masculine appearance. This shows that while some words have the same core meaning, their usage characteristics may vary.

Chung (2011) examined the similarities and distinctions between the synonyms *'create'* and *'produce'* using corpus-based analysis. Using the Sketch Engine (SkE), data from the Brown Corpus and the Freiburg-Brown corpus of American English (FROWN) were compared to the British National Corpus (BNC). Both words share two overlapping meanings and are most frequently used in their infinitive and -ed forms, according to the study. The overlapping meanings are *'bring*

into existence/cause to happen, occur, or exist' and 'create or manufacture a man-made product.' However, the objects following each verb are distinct. *'Produce'* is typically followed by objects manufactured in large quantities with low creativity, such as factory-made goods; additionally, the objects following *'produce'* are diverse, including crops and manufactured goods, due to its reference to factory-made products. *'Create'* however, is usually followed by fewer abstract objects whose properties are not fixed, such as problems or images, that allow for greater creativity. The study illuminates the subtle differences in meaning and usage between synonyms and emphasizes the significance of corpus-based analysis for examining language usage.

Using data from the Corpus of Contemporary American English (COCA), Cai conducted a study on seven near-synonyms in 2012: *'awesome,' 'excellent,' 'fabulous,' 'fantastic,' 'great,' 'tremendous,'* and *'wonderful'*. The purpose of the study was to identify any variations in denotation or connotation, collocational usage, and frequency distribution among genres. The study revealed that *'great'* was the most frequently occurring near-synonym and also had a broader meaning. Furthermore, *'fabulous,' 'fantastic,' 'great,' 'terrific,'* and *'wonderful'* were more prevalent in spoken language, whereas *'awesome'* and *'excellent'* were more prevalent in magazines. The number of collocates of the target synonyms was found to vary considerably, but they shared some noun collocates. The majority of near-synonyms were used to modify abstract nouns to describe concepts and ideas, with *'wonderful'* being the most common, whereas *'excellent'* was more frequently used to describe movement and events. In addition, the study uncovered differences in connotation among the near-synonyms. *'Great'* was viewed as more neutral than the other terms, whereas *'fabulous,' 'fantastic,' 'terrific,'* and *'awesome'* were viewed as highly positive. Both *'fabulous'* and *'fantastic'* had an almost unbelievable quality. The conclusion of the study was that there are typically differences in the meanings and applications of apparent synonyms and that lexical items are rarely interchangeable.

Jarunwaraphan and Mallikamas (2020) used the Corpus of Contemporary American English (COCA) to compare *'chance'* and *'opportunity,'* two English synonyms. They selected these nouns due to their variety of meanings and analyzed their definitions, idioms, expressions, and grammatical patterns using online dictionaries, including Longman Dictionary of Contemporary English, Cambridge Dictionary, and Merriam-Webster Dictionary. The study revealed that the two words are not interchangeable and that their usage varies based on context and collocation. *'Opportunity'* is typically associated with positive connotations such as *'create,' 'expand,'* and *'promote,'* whereas *'chance'* is frequently associated with negative connotations such as *'damage,' 'eliminate,'* and *'reduce'*. In terms of formality, the study revealed that *'opportunity'* is more frequent in academic texts, while *'chance'* is more noticeable in spoken language. The study concluded that the two synonyms have distinct distributions across genres and that *'chance'* has a wider range of meanings than *'opportunity'*. Instead of focusing solely on grammaticality, the authors suggested that vocabulary should be taught in context with precise examples of usage and that corpus data could demonstrate how a grammatical pattern can be used in different contexts.

By using diverse criteria and resources, such as corpora and dictionaries, numerous studies have demonstrated how to differentiate synonyms. Recent years have seen an increase in the popularity of corpus-based studies, particularly in the analysis of collocations through frequency and mutual information (MI) scores (Alanazi, 2023; Imsa-ard, 2021; Lertcharoenwanich, 2023; Phoocharoensil, 2022; Sittironnarit et al., 2022). This study was motivated by such an analysis to differentiate the synonyms *'important,' 'significant,'* and *'crucial,'* focusing on their distribution across eight genres in the Corpus of Contemporary American English (COCA), as there do not appear to be any corpus-based studies investigating these synonyms in terms of their distribution across genres and collocations. To examine these adjectives in depth, the study employed the criteria of frequency, collocations, degree of formality, and semantic preference. The objective of the research was to provide empirical evidence on the corpus-based study of synonyms and to expand available linguistic data that may enhance EFL learning. Moreover, the findings of this study could be applied to English language instruction in order to help students who may have trouble using synonyms and to reduce any misunderstandings that could result in communication breakdown

(Jarunwaraphan & Mallikamas, 2020). The following research questions are addressed to achieve the aim of the study:

1. How do the synonyms '*important*', '*significant*', and '*crucial*' differ across different genres?
2. What are the common collocations of the synonymous adjectives '*important*', '*significant*', and '*crucial*'?

Methodology

Data Collection & Data Analysis

This study examined the synonymous adjectives '*important*,' '*significant*,' and '*crucial*' due to their high frequency of use and inherent significance in the realm of language acquisition. To acquire their primary meanings, definitions, example sentences, and usage notes, three online American-English dictionaries were consulted, specifically Longman Advanced American Dictionary, Oxford Advanced American Dictionary, and Merriam-Webster Dictionary. Then, the Corpus of Contemporary American English (COCA) was utilized as the primary instrument to determine the frequency of the three synonyms across genres and noun and adverb collocates. COCA is a huge compilation of texts that contains more than one billion words of text and has been updated annually from 1990 to 2022, with approximately 25 million words added each year (Davies, 2020). The corpus is unique in that it includes a wide variety of genres, such as spoken language, fiction, popular magazines, newspapers, academic texts, TV and movie subtitles, blogs, and online web pages. This allows researchers to examine language usage in a variety of contexts and investigate how language varies across genres. A key characteristic of COCA is its size. This makes it an indispensable resource for analyzing language use in various contexts and identifying patterns and trends in language use over time. Another essential aspect of COCA is its balance between genres. The corpus is designed to include roughly equal amounts of data from each genre, allowing researchers to compare language usage across genres and search for differences and similarities in the way language is used in various contexts. Lastly, COCA is regularly updated with new data, allowing researchers to study language usage in real time and track changes over time. In corpus-based research, COCA provides authentic American English data that facilitates more practical and systematic analyses of genre as well as linguistic and lexical features (Alanazi, 2023; Bergdahl, 2009; Phoocharoensil, 2020a; Phoocharoensil, 2022). This makes COCA a valuable resource not only for studying language use but also for comprehending how language use has evolved over the past several decades.

The current study addresses two research questions. Using the Corpus of Contemporary American English (COCA), the first research question investigated the distribution and frequency of the target synonyms, namely '*important*,' '*significant*,' and '*crucial*,' across eight different genres. In order to answer the second research question, the researcher examined the frequent noun and adverb collocates of the three target synonyms to determine their semantic preference. The strength of collocation was measured using the Mutual Information (MI) value (Gablasova et al., 2017), which determines whether two words co-occur by chance or have a strong collocational association. However, the MI value has limitations in that it cannot provide precise results of collocational strength because it may provide a high value for combinations with low frequency (Cheng, 2012). That is, collocation research that relies solely on MI scores may place a greater emphasis on uncommon combinations, as the MI value tends to favor these words when ranking the collocates. To ensure that only frequent, recurrent collocations were extracted, the researcher measured the strength of collocation in this study by combining frequency and the MI value. The top 30 frequency lists presented in COCA with an MI score value that meets the significance level for collocational association of 3 or higher (Smyth 2016), were chosen for the selection of noun collocates. The same criteria were applied to the top-20 frequency list of adverb collocates. To elaborate, collocations that are semantically related were grouped together under the same theme.

This semantically related categorization provided a better understanding of the differences in collocational patterns and context of occurrence among the target synonyms.

Results and Discussion

This section presents the findings of an investigation that aimed to answer two research questions by analyzing the frequency and distribution of three synonyms across different genres and then examining their noun and adverb collocations. A discussion of the principal findings concludes the section.

Frequency and Distribution of the Synonyms across Genres

The findings with regard to the distribution of '*important*,' '*significant*,' and '*crucial*' across genres in COCA will be shown and discussed.

Table 2

Distribution of '*important*,' '*significant*,' and '*crucial*' across eight genres according to frequency

<i>important</i>			<i>significant</i>			<i>crucial</i>		
Genre	Freq.	Per million	Genre	Freq.	Per million	Genre	Freq.	Per million
academic texts	76,605	639.49	academic texts	60,470	504.80	academic texts	7,965	66.49
spoken	51,467	408.03	webpages	15,190	122.25	magazines	5,338	42.33
webpages	46,725	376.05	blogs	13,194	102.59	newspaper	4,142	34.02
blogs	46,809	363.95	magazines	11,590	91.92	webpages	3,508	28.23
magazines	41,134	326.22	newspaper	10,700	87.89	blogs	3,151	24.50
newspaper	29,000	238.21	spoken	9,002	71.37	spoken	2,504	19.85
TV and movies subtitles	25,290	197.46	fiction	1,711	14.46	fiction	890	7.52
fiction	14,649	123.81	TV and movies subtitles	1,247	9.74	TV and movies subtitles	748	5.84
Total	331,679		Total	123,104		Total	28,246	

According to Table 2, '*important*' is the most prevalent of the three near-synonyms. In particular, the target adjective '*important*' appeared almost three times more frequently than '*significant*' (123,104 tokens). The adjective '*crucial*' seems to be used the least (28,246 tokens). Regarding genre, it is evident that the near-synonyms '*important*,' '*significant*,' and '*crucial*' are generated more frequently in formal than informal contexts. As demonstrated in Table 2 by the frequency of occurrences, the lexical items of focus are predominantly found in academic texts. Thus, '*important*,' with 76,605 tokens, is the most frequent of the three targets in academic texts and indicates the most formal language. The adjective '*significant*' occurs the second most frequently with 60,470 tokens, while '*crucial*' occurs the least frequently with 7,965 tokens.

In addition, Table 2 reveals that the distribution of the target synonyms is comparable. Evidently, the terms '*important*,' '*significant*,' and '*crucial*' occur least frequently in informal contexts.

In particular, *'crucial'* appears with the lowest frequency in the three genres of TV and movies subtitles (748 tokens), fiction (890 tokens), and spoken (2,504 tokens) that contain its synonyms. In terms of informality, the distribution of *'significant'* and *'crucial'* is identical. *'Significant'* is used the least in TV and movie subtitles (1,247 tokens), fiction (1,711 tokens), and spoken language (9,002 tokens).

Table 2 also shows that the three synonymous adjectives *'important,' 'significant,'* and *'crucial'* are frequently found in web pages, which is a new genre embedded within COCA. The frequent occurrences of these synonyms suggest that web pages contain both academic and non-academic texts (e.g., Phoocharoensil, 2020a; Phoocharoensil, 2020b). These findings are consistent with previous studies (Cai, 2012; Petcharat & Phoocharoensil, 2017) and suggest that, despite their varying occurrences across genres, the three synonyms have comparable levels of formality.

Common collocations: Analysis of noun and adverb collocates

The following section answers the second research question by analyzing the noun and adverb collocates of the synonymous adjectives *'important,' 'significant,'* and *'crucial'*. Specifically, the analysis focuses on frequently occurring collocates with an MI value of 3 or above and a frequency higher than 5. Additionally, the semantic preference of the collocates is also discussed.

Noun Collocates and Semantic Preference of *'important,' 'significant,'* and *'crucial'*

Table 3

Noun Collocates of *'important,' 'significant,'* and *'crucial'* in COCA

Rank	<i>important</i>			<i>significant</i>			<i>crucial</i>		
	Noun collocate	Freq.	MI Value	Noun collocate	Freq.	MI Value	Noun collocate	Freq.	MI Value
1	role	6671	3.76	difference	11513	5.97	role	1428	5.09
2	factor	4339	3.71	effect	4632	4.42	moment	523	3.22
3	aspect	2401	3.81	change	3626	3.74	element	461	4.67
4	lesson	1487	3.07	amount	2456	4.27	step	393	3.51
5	component	1315	3.39	impact	2104	4.33	factor	376	3.74
6	implication	1101	3.74	role	2009	3.46	difference	354	3.07
7	consideration	1087	3.68	increase	1786	4.66	component	246	4.53
8	contribution	1034	3.15	relationship	1756	3.21	aspect	217	3.90
9	distinction	902	3.88	correlation	1740	6.32	vote	205	3.00
10	determinant	316	5.17	improvement	1559	5.20	importance	188	4.01
11	predictor	267	3.53	interaction	1498	5.11	distinction	144	4.79
12	milestone	261	4.30	factor	1483	3.59	understanding	143	3.11
13	caveat	150	4.03	contribution	1326	4.93	survival	98	3.95
14	constituency	126	3.01	portion	1247	5.12	insight	75	3.41
15	facet	114	3.31	predictor	1241	7.18	junction	74	7.29
16	takeaway	61	3.41	reduction	1057	5.01	ingredient	70	3.58
17	correlate	31	3.25	variable	996	4.47	battleground	45	5.87
18	personage	18	3.34	association	798	3.10	primary	40	3.19
19	confounder	10	3.29	progress	709	3.67	habitat	33	3.08
20				finding	702	3.57	determinant	25	5.07
21				influence	575	3.29	interception	14	3.18
22				damage	563	3.23	crossroads	9	3.35
23				decrease	554	6.01	caveat	9	3.53
24				gain	522	3.96			
25				percentage	498	3.41			
26				variance	480	5.43			
27				proportion	465	4.41			
28				shift	430	3.61			
29				achievement	397	3.44			
30				implication	367	3.58			

Table 3 displays the top-30 noun collocates, but based on the frequency and MI score criteria (MI value of 3 at the minimum), there are only 19 noun collocates of *'important'* and 23 noun collates of *'crucial'*. In this study, the three target adjectives are considered near-synonyms because they share some common noun collocates, which suggests that they are synonyms (Phoocharoensil, 2020a). This is evident from the table presented.

The analysis of the noun collocates of the three nearly synonymous adjectives *'important,'* *'significant,'* and *'crucial'* reveals that the noun collocates *role* and *factor* are shared by all three. There are more nouns that frequently co-occur with *'important'* and *'significant,'* such as *implication,* *contribution,* and *predictor*. However, corpus data suggests that *'important'* and *'crucial'* are close synonyms because they share several noun collocates such as *aspect,* *component,* *determinant,* *distinction,* and *caveat*. Other than *difference,* there are no similar noun collocates between *'significant'* and *'crucial'*. This suggests a closer relationship between *'important'* and *'crucial'*. This finding is consistent with previous studies (Phoocharoensil, 2020a; Sinclair, 2004; Stubbs, 2001), indicating that words with similar meanings tend to share collocates.

Although Table 3 contains information on collocations, it is essential to note that only the top 30 nouns were examined. Due to their low frequency or MI score in COCA, there may be additional nouns that can co-occur with the three target adjectives but were not included in the lists above. When interpreting the results, it is necessary to keep this limitation in mind. Nonetheless, the collocational data presented in Table 3 is worth examining in depth. The data can still provide valuable insights despite the absence of some nouns associated with *'important,'* *'significant,'* and *'crucial'* due to their low frequency or MI score in COCA. For example, despite the high frequency of the noun collocation *'important thing'* in COCA (16,964 tokens), it was excluded from the list due to its MI score of 2.40. Similarly, the number of occurrences of *'crucial part'* in COCA (591 tokens) is also high, but its MI score of 2.33 led to its exclusion from the list.

There are frequent co-occurrences of *'important'* and research-related matters such as *determinant,* *predictor,* *correlate,* *confounder,* and *implication*. The near-synonyms *'important,'* *'significant,'* and *'crucial'* have shared and different noun collocates, indicating that their usage is heavily dependent on semantic preference. In the next section, the 30 most frequent noun collocations of each target synonym were categorized according to their semantic preference in order to determine similarities in meaning between the collocations (Phoocharoensil, 2020).

Table 4

Semantic Preference of Noun Collocates of *'Important'*

1. RESEARCH-RELATED MATTERS	<i>confounder, correlate, determinant, implication, predictor</i>
2. TAKEAWAY	<i>lesson, milestone, takeaway</i>
3. ROLES	<i>personage, role</i>
4. CAUSE	<i>contribution, factor</i>
5. PERSPECTIVE	<i>aspect, facet</i>
6. COMPONENTS	<i>component</i>
7. DEVIATION / FLUCTUATION	<i>distinction</i>
8. DECISION-MAKING	<i>consideration</i>
9. POLITICAL DOMAIN	<i>constituency</i>
10. MISCELLANEOUS	<i>caveat</i>

The noun collocates for the word *'important'* are divided into nine major themes in Table 4: RESEARCH-RELATED MATTERS, TAKEAWAY, ROLES, CAUSE, PERSPECTIVE, COMPONENTS, DEVIATION/FLUCTUATION, DECISION-MAKING, and POLITICAL DOMAIN. These themes reflect the various semantic relationships that *'important'* can have with its collocates and demonstrate the versatility of the term in various contexts. Several examples of concordance lines from COCA are provided to aid comprehension of collocation behaviors in context. As demonstrated in example (3), the majority of the noun collocates of the adjective

'important' fall under the category of RESEARCH-RELATED MATTERS, i.e., *confounder*, *correlate*, *determinant*, *implication*, and *predictor*, which are all relevant to 'the research being conducted.' The second most frequent theme, TAKEAWAY, consists of nouns indicating 'what can be learned or deduced from a research study.' As shown in (4), this theme includes the terms *lesson*, *milestone*, and *takeaway*. It is noteworthy that two themes are closely related to the adjective 'important'. These are ROLES, such as *personage* and *role*, as shown in example (5), and CAUSE, including *contribution* and *factor*, as illustrated in example (6). For the PERSPECTIVE theme, 'important' also co-occurs with *aspect* and *facet*, which convey 'different angles, viewpoints, or dimensions in doing things' as exemplified in (7). The sixth theme classifies one noun, *component*, into the theme COMPONENTS, as shown in (8). For the DEVIATION/FLUCTUATION theme, there is one co-occurring noun, *distinction*, that highlights 'differences or changes in a specific situation or phenomenon,' as shown in (9). For the last two main themes, *consideration* is classified into the DECISION-MAKING theme, as shown in (10), and *constituency* is put under the theme POLITICAL DOMAIN as it relates to politics and governance, as illustrated in (11).

- (3) Studies find that the most ***important determinant*** of an organization's climate is the day-to-day behavior of its leaders.
- (4) Cultural innovations implemented at the time were ***important milestones*** to secure the active roles of women as individuals having free will and free...
- (5) Laboratory and imaging findings have an ***important role*** in identifying patients with transmural necrosis who might benefit from emergency surgical treatment.
- (6) Plant species identity can be an ***important factor*** influencing root microbial diversity; however, this relationship is not universal.
- (7) Media content is an ***important aspect*** of that reaffirmation.
- (8) Time is an ***important component*** of organizational effectiveness and strategic planning.
- (9) This is an ***important distinction*** that must be made clear when discussing legal liabilities.
- (10) Another ***important consideration*** or potential limitation is that of the power differentials.
- (11) There is basically no ***important constituency*** that has been left out of the loop, and that is quite deliberate...

Table 5

Semantic Preference of Noun Collocates of 'Significant'

1. RESEARCH-RELATED MATTERS	<i>correlation, finding, implication, predictor, variable, variance</i>
2. ADVANCEMENT	<i>achievement, gain, improvement, increase, progress</i>
3. CONNECTION	<i>association, interaction, relationship</i>
4. QUANTITY	<i>amount, percentage, portion, proportion</i>
5. CAUSE	<i>contribution, factor, influence</i>
6. DEVIATION / FLUCTUATION	<i>change, difference, shift</i>
7. DECLINE	<i>damage, decrease, reduction</i>
8. OUTCOME	<i>effect, impact</i>
9. ROLES	<i>role</i>

Table 5 displays the classification of 'significant' noun collocates into nine themes. These themes were also identified through an examination of the word's semantic preferences. Three of the themes associated with the adjective 'significant' are observed to overlap with those that categorize 'important.' These themes include RESEARCH-RELATED MATTERS, which includes

terms such as *correlation*, *finding*, *implication*, *predictor*, *variable*, and *variance*, as shown in (12), CAUSE, which includes terms such as *contribution*, *factor*, and *influence*, as in (13), and ROLES, which includes the term ‘*role*’, as in (14).

- (12) The fact that the rehabilitation system is based mainly on factory operations has a **significant implication**.
- (13) The intimate account remains a **significant contribution** to Native American studies, literary studies, and anthropology.
- (14) Targets of miRNA have a **significant role** in viral pathogenesis and host defense mechanisms.

The themes ADVANCEMENT and CONNECTION have a considerable number of noun collocates, ranked second and third, respectively. ADVANCEMENT includes nouns that suggest ‘the idea of improvement or progress,’ such as *achievement*, *gain*, *improvement*, *increase*, and *progress*, as shown in (15). Likewise, CONNECTION, which includes *association*, *interaction*, and *relationship*, as shown in (16), should be highlighted.

- (15) Early intervention during the preschool years may yield **significant improvement** in reading and math skills for those children at greater risk of academic difficulties.
- (16) The technological intensity of the industry has a **significant relationship** with the proportion of output that is exported.

Although the DEVIATION/FLUCTUATION theme is also part of the adjective ‘*crucial*’, as presented in Table 6 below, there are three other themes that are solely connected to ‘*significant*,’ namely QUANTITY, DECLINE, and OUTCOME. Quantity-related noun collocates include *amount*, *percentage*, *portion*, and *proportion*, as shown in (17), whereas DECLINE-related noun collocates include *damage*, *decline*, and *reduction*, as shown in (18). It is important to note that the OUTCOME theme consists of only two members, which are *effect* and *impact*, as depicted in (19).

- (17) The reality is that a **significant proportion** of these adolescents demonstrate substantial need for prevention and intervention services.
- (18) All districts experienced a **significant decline** in crop farming as a source of income.
- (19) The diversity in our workshop had a **significant impact** on the process of learning.

Table 6

Semantic Preference of Noun Collocates of ‘*Crucial*’

1. POLITICAL DOMAIN	<i>battleground, primary, survival, vote</i>
2. COMPONENTS	<i>component, element, ingredient</i>
3. ADVANCEMENT	<i>insight, step, understanding</i>
4. DEVIATION / FLUCTUATION	<i>difference, distinction</i>
5. TIME	<i>juncture, moment</i>
6. ROLES	<i>role</i>
7. CAUSE	<i>factor</i>
8. RESEARCH-RELATED MATTERS	<i>determinant</i>
9. PERSPECTIVE	<i>aspect</i>
10. ENVIRONMENTAL HEALTH	<i>habitat</i>
11. MISCELLANEOUS	<i>caveat, crossroads, importance, interception</i>

Table 6 shows that two themes related to both '*important*' and '*significant*' were found for the adjective '*crucial*': ROLES and CAUSE. These themes are typically associated with the nouns *role*, as in (20), and *factor*, as in (21), respectively. In addition, the noun *component*, presented in (22), is a common collocate of '*important*,' and *element* and *ingredient* are two more collocates of '*crucial*,' as exemplified in (23). In addition, the noun *aspect*, which falls under the theme PERSPECTIVE, is also a common collocate of '*important*,' as shown in (24).

- (20) Encouraging university teachers to work on an innovative task together does play a ***crucial role*** in achieving educational change.
- (21) Measuring intervention sustainability is a ***crucial factor*** in maintaining a community's health and well-being.
- (22) Geographers have shown, however, that it is a ***crucial component*** of capitalism to unevenly exploit space, investing in certain places and disinvesting in others.
- (23) Creating interaction with students is a ***crucial element*** of establishing pace and could have greatly helped flow of the lecture and offset...
- (24) Without a doubt, individualism is a ***crucial aspect*** of the social imaginary of modernity, providing real individuals with an understanding of who...

As demonstrated in (25), the POLITICAL DOMAIN theme has the greatest number of noun collocates among all the themes, including *battleground*, *primary*, *survival*, and *vote*. In addition, the noun *difference*, as presented in (26) under the DEVIATION/FLUCTUATION theme, is frequently associated with '*significant*,' whereas the noun *distinction*, as demonstrated in (27), is a common collocate of '*important*.' The TIME theme, which includes *juncture* and *moment*, as in (28), and the ENVIRONMENTAL HEALTH theme, which includes *habitat*, as in (29), are unrelated to the other two synonyms. This discovery makes it significantly easier to differentiate between these three similar terms.

- (25) Prime Minister Theresa May's prospects of winning the ***crucial vote*** were dealt a significant blow Tuesday morning when the attorney general, Geoffrey Cox...
- (26) What teachers know and can do makes the ***crucial difference*** in what children learn.
- (27) An anthropologist by training, makes a ***crucial distinction*** between two definitions of diversity.
- (28) This innovation marks a ***crucial moment*** in the development of the pictorial mode known as "genre."
- (29) In addition to the services provided, riparian forests provide ***crucial habitat*** for the conservation of biodiversity at multiple scales.

The analysis of the noun collocates '*important*,' '*significant*,' and '*crucial*' reveals both their usage similarities and differences. These synonyms are most prevalent in formal written English, with the highest frequency in academic texts. All three adjectives share two noun collocates, *role*, which is under the theme ROLES, and *factor*, which is under the theme CAUSE, indicating a semantic relationship. Moreover, another theme that '*important*,' '*significant*,' and '*crucial*' all share is the theme of RESEARCH-RELATED MATTERS. The noun collocates in this theme with research studies, statistical analysis, and modeling. However, by analyzing the nouns that frequently appear alongside each adjective, we can gain insight into their distinctive qualities. The adjective '*important*'

is far more prevalent in spoken than ‘*significant*’ and ‘*crucial*’. Interestingly, the frequency of ‘*important*’ is lowest in fiction, whereas ‘*significant*’ and ‘*crucial*’ are the least common in TV/movie subtitles. These findings are consistent with previous studies by Phoocharoensil (2020a), Sinclair (2004), and Stubbs (2001), which indicate that certain synonyms are more frequently used in particular genres.

Additionally, this study not only analyzed the frequency of the three synonyms of adjective across different genres in COCA, but also their collocational patterns. By classifying typical collocates into themes, this analysis reveals both the similarities and differences in how these synonyms are used in a more specific context. The themes of PERSPECTIVE and COMPONENTS are shared by ‘*important*’ and ‘*significant*’, whereas the theme of DEVIATION/FLUCTUATION is shared by ‘*significant*’ and ‘*crucial*’. The characteristics of the noun collocates of each synonymous adjective were differentiated upon close examination. For instance, some nouns appear to be particularly combined with ‘*important*’ in relating to TAKEAWAY, e.g., *lesson*, *milestone*, and *takeaway*, and DECISION-MAKING, e.g., *consideration*, whereas some nouns exclusively co-exist with ‘*significant*’, as opposed to ‘*important*’ and ‘*crucial*’, relating to QUANTITY, e.g., *amount*, *percentage*, *portion*, and *proportion*, CONNECTION, e.g., *association*, *correlation*, *interaction*, and *relationship*, DECLINE, e.g., *damage*, *decrease*, and *reduction*, and OUTCOME, e.g., *effect* and *impact*. This suggests that the three target words are only slightly synonymous, and their status as near-synonymous is weak. Regarding this particular point, previous research has provided support for the differences in usage and semantic preference of each synonym based on the distinctions in collocational patterns (Cheng, 2012; Ly & Jung 2015; Phoocharoensil, 2022; Yang, 2016).

Adverb Collocates and Semantic Preference of ‘*important*,’ ‘*significant*,’ and ‘*crucial*’

Table 7

Adverb Collocates of ‘*important*,’ ‘*significant*,’ and ‘*crucial*’ in COCA

Rank	<i>important</i>			<i>significant</i>			<i>crucial</i>		
	Adverb collocate	Freq .	MI Value	Adverb collocate	Freq .	MI Value	Adverb collocate	Fr eq.	MI Value
1	extremely	2468	4.16	statistically	6731	9.26	absolutely	199	3.15
2	equally	2268	4.78	highly	656	3.27	strategically	10	3.72
3	increasingly	1197	3.32	clinically	349	7.04	hotly	3	3.08
4	critically	1051	5.74	potentially	273	3.17	vitality	2	3.04
5	vitality	824	8.17	historically	237	3.93	stunningly	2	3.10
6	incredibly	649	3.39	marginally	182	6.41	journalistically	2	6.17
7	terribly	378	3.65	culturally	113	3.85			
8	arguably	327	3.82	hugely	60	3.99			
9	hugely	325	5.00	pairwise	53	6.10			
10	strategically	216	4.60	architecturally	33	6.32			
11	enormously	199	4.13	biologically	27	3.60			
12	crucially	172	5.83	ecologically	24	4.26			
13	tremendously	130	3.64	symbolically	22	3.79			
14	extraordinarily	129	3.32	evolutionarily	20	5.79			
15	clinically	113	3.98	militarily	19	3.09			
16	immensely	91	3.30	educationally	17	4.83			
17	symbolically	89	4.37	theologically	15	4.08			
18	systemically	65	6.14	ritually	15	5.21			
19	biologically	52	3.11	regionally	10	3.53			
20	supremely	50	3.90	systemically	10	4.87			

Only the top-20 adverb collocates are shown in Table 7, but based on the frequency and MI score criteria (minimum MI value of 3) there are only six adverb collocates of ‘*crucial*’. The table reveals that three synonymous adjectives share adverb collocates. In particular, ‘*important*’ and

'significant' share adverb collocates such as *clinically*, *biologically*, and *systemically*, which are associated with scientific analysis or methodical approaches. Another shared adverb collocate is *symbolically*, which relates to the ideas of significance and relevance. In addition, the adjectives 'important' and 'crucial' share the collocates *vitally* and *strategically*. However, 'significant' and 'crucial' do not share any adverb collocates on the list of the 20 most frequent adverbs from COCA. Although there are a number of common adverb collocates between 'important' and 'significant,' there are also distinct collocates that help to differentiate them to some degree. There may be additional adverb collocates shared by these three synonyms, but those with an MI score below 3 were omitted, resulting in their possible absence from Table 7. Although common adverbs such as *most* can be used along with the target words, such as *most important*, *most significant*, and *most crucial*, these weaker collocations were not included on the list. This decision was made as a result of the fact that these collocations pose less of a challenge for students and, consequently, are not considered technically valuable (Phoocharoensil, 2022).

The following section provides the semantic preference of adverb collocates for the three synonyms of the adjective, based on their similarity in meaning. Understanding collocations may be helpful in determining the range of associations between terms and identifying the semantic relationships between their collocates, according to previous research by Ly and Jung (2015).

Table 8

Semantic Preference of Adverb Collocates of 'Important'

1. INTENSITY	<i>enormously, extraordinarily, extremely, hugely, immensely, increasingly, incredibly, supremely, terribly, tremendously</i>
2. IMPORTANCE	<i>arguably, critically, crucially, equally, strategically, symbolically, vitally</i>
3. SCIENTIFIC ANALYSIS / METHODICAL	<i>biologically, clinically, systemically</i>

As shown in Table 8, three major themes emerged from the semantic preference analysis of 'important,' namely INTENSITY, IMPORTANCE, and SCIENTIFIC ANALYSIS / METHODICAL. The adverb collocates under the theme of INTENSITY account for the majority of the collocates. The collocates under this theme include adverbs indicating a significant or heightened level of intensity or magnitude, i.e., *enormously*, *extraordinarily*, *extremely*, *hugely*, *immensely*, *increasingly*, *incredibly*, *supremely*, *terribly*, and *tremendously*, as exemplified in (30). The second most frequent theme, IMPORTANCE, consists of adverb collocates such as *arguably*, *critically*, *crucially*, *equally*, *strategically*, *symbolically*, and *vitality*, as in (31). Notably, the adverb *symbolically*, which is related to the adjective 'significant,' overlaps with the collocates that categorize 'important.' Furthermore, the adverb collocates *strategically* and *vitality* are also found to coexist with the adjective 'crucial,' as presented in (32).

- (30) During the early mobilization period this group was an ***enormously important*** politico-economic group, thrashing out top policy, advising major governmental agencies.
- (31) In the case of cash management, it is ***arguably important*** to audit cash positions frequently, and, at the same time...
- (32) Some industries, such as oil, are considered ***strategically important*** to the regime, whereas others, such as telecommunications, are politically...

Table 9

Semantic Preference of Adverb Collocates of 'Significant'

1. FIELDS OF STUDY	<i>architecturally, ecologically, educationally, historically, militarily, theologically</i>
2. SCIENTIFIC ANALYSIS / METHODICAL	<i>biologically, clinically, evolutionarily, statistically, systemically</i>
3. INTENSITY	<i>highly, hugely, marginally</i>
4. IMPORTANCE	<i>potentially, ritually, symbolically</i>
5. CULTURE	<i>culturally</i>
6. MISCELLANEOUS	<i>pairwise, regionally</i>

Table 9 illustrates the classification of 'significant' adverb collocates into five different themes: FIELDS OF STUDY, SCIENTIFIC ANALYSIS / METHODICAL, INTENSITY, IMPORTANCE, and CULTURE. The majority of these adverbs collocates fall under the FIELDS OF STUDY theme, which includes adverbs associated with specific areas of study, such as *architecturally, ecologically, educationally, historically, militarily, and theologically*, as exemplified in (33). The following items present selected instances from the concordance lines that exemplify the overlapping themes of SCIENTIFIC ANALYSIS / METHODICAL, INTENSITY, and IMPORTANCE with the 'important' themes. These themes encompass terms that highlight specific aspects. For SCIENTIFIC ANALYSIS / METHODICAL, notable examples include *biologically, clinically, evolutionarily, statistically, and systemically*, as seen in (34). For IMPORTANCE, the terms *potentially, ritually, and symbolically* are evident, as shown in (35). As for INTENSITY, examples include *highly, hugely, and marginally*, as demonstrated in (36).

- (33) These are changes that are ***educationally significant*** since they represent building blocks for future learning.
- (34) In this study, mental illness refers to ***clinically significant*** disturbances affecting the individual's relationship patterns.
- (35) This distribution of packages of pork and rice is ***symbolically significant*** in at least two ways.
- (36) Analysis of variance performed for genotypes showed the existence of ***highly significant*** differences between the three genotypes.

The presence of the CULTURE theme, represented by the adverb *culturally*, as shown in (37), is distinct from the other two synonyms and of significant importance. This finding greatly facilitates the differentiation of these three synonymous terms. Two adverbs, *pairwise* and *regionally*, do not appear to relate to any of the established themes; therefore, they have been classified as MISCELLANEOUS. However, it is important to note that new themes may emerge as future research incorporates a broader range of potential adverb collocates. This would allow the miscellaneous collocates to be reassigned to these emerging themes.

- (37) Motherhood was a secondary role for women, less ***culturally significant*** than women's work as economic providers.

Table 10

Semantic Preference of Adverb Collocates of 'Crucial'

1. INTENSITY	<i>absolutely, hotly, stunningly</i>
2. IMPORTANCE	<i>strategically, vitally</i>
3. FIELDS OF STUDY	<i>journalistically</i>

After conducting a thorough analysis of the semantic preference of '*crucial*' adverb collocates, three primary themes were identified, as shown in Table 10. These themes, INTENSITY, IMPORTANCE, and FIELDS OF STUDY, are consistent with those observed for the near-synonyms '*important*' in the INTENSITY and IMPORTANCE themes, as well as '*significant*' across all of those themes.

- (38) The next ***absolutely crucial*** linkage is that the services must be demonstrated to produce the desired results.

By extracting and categorizing typical collocations into various themes, it is possible to observe both similarities and differences in the collocational behavior of these synonyms with more completeness. As with noun collocates, the results of adverb collocates provided insight into the usage distinctions between each synonymous adjective. Despite the fact that they share some collocates, it is evident that these collocates represent only a part of their overall usage. The themes of INTENSITY and IMPORTANCE are shared by the frequent adverb collocates '*important*,' '*significant*,' and '*crucial*'. However, a closer look reveals that individual occurrences of the same theme may vary. For instance, while '*important*' and '*significant*' frequently co-occur with adverbs such as *biologically*, *clinically*, *systemically*, and *symbolically*, the adverbs *strategically* and *vitality* form stronger collocations with '*important*' and '*crucial*' than with '*significant*.' In addition, the adjective '*significant*' has distinct adverb collocates associated with CULTURE, such as *culturally*. This indicates that these synonyms are not identical with regard to collocation. These variations in typical collocations shed light on the usage differences between the three synonyms, which would be difficult to explain without corpus-based data (e.g., Chung, 2011; Nisani, 2015; Phoocharoensil, 2020a).

The results of the study indicate that using corpus-based analysis to study synonyms offers more insightful information than relying solely on dictionaries (Aroonmanakun, 2015). Consistent with previous research (Phoocharoensil, 2010; Jarunwaraphan and Mallikamas, 2020; Thornbury, 2002), the results of this study suggest that the usage of some synonymous words varies in different contexts. In addition, the adjectives '*important*,' '*significant*,' and '*crucial*' are categorized as near-synonyms, and while they share some similarities, they have different semantic properties and collocations (Aroonmanakun, 2015; Jarunwaraphan & Mallikamas, 2020; Phoocharoensil, 2020; Stubbs, 2001; Taylor, 2003). Overall, these three adjectives share some collocations with nouns and adverbs, but their unique collocations distinguish them from one another.

Conclusion

This corpus-based research aimed to investigate the distribution of three near-synonyms, namely '*important*,' '*significant*,' and '*crucial*,' across different genres and their frequent noun and adverb collocates. The study demonstrates the similarities and differences among these words, revealing that they share a core meaning but differ in their detailed meanings and collocation patterns. By drawing data from dictionaries and the Corpus of Contemporary American English (COCA), the study highlights how corpora provide additional information not found solely in dictionaries.

As the study examines the distribution and collocation of the synonyms '*important*,' '*significant*,' and '*crucial*' using COCA data, these adjectives are prevalent in academic writing, with '*important*' being the most commonly used. As for informality, however, there is no distinction between '*significant*' and '*crucial*.' While they share some collocation patterns, distinct noun and adverb collocations are also identified, indicating that each adjective has its own unique semantic associations. For instance, '*significant*' is associated with quantity-related and downward trend concepts, whereas '*crucial*' is associated with the political domain. The analysis of collocates between nouns and adverbs reveals both shared and distinct preferences among synonymous adjectives. '*Important*' frequently co-occurs with adverbs of intensity; '*significant*' frequently co-

occurs with adverbs related to fields of study; and *'crucial'* demonstrates theme of culture that is unique to it. In addition to confirming the synonymy of the adjectives, these results highlight the subtle differences in their collocational patterns.

In this study, formality and collocation play a crucial role in distinguishing the synonyms *'important,' 'significant,'* and *'crucial'* from each other. Due to these distinctions, the adjectives cannot always be used interchangeably. It should be noted that corpora provide additional information beyond what is available in dictionaries. Furthermore, the extensive collocational data derived from corpora like COCA proves valuable language teaching materials for educators (Phoocharoensil, 2020a). Educators should guide students in understanding the distinctions between synonyms and their appropriate usage. Consequently, using corpus methods to illustrate genre distinctions and formality serves as a crucial and effective criterion for differentiating synonyms.

This study has some limitations that need to be addressed. The investigation focuses mainly on three synonyms: *'important,' 'significant,'* and *'crucial.'* Future research should investigate additional terms from the synonym collection, such as *'essential,' 'necessary,'* and *'vital,'* in order to achieve a more comprehensive understanding. In addition, the analysis relies only on the Corpus of Contemporary American English (COCA) data and does not include a comparison between American English and British English using the British National Corpus (BNC). Incorporating multiple corpora into future studies would therefore provide greater insight into the distribution of these near-synonyms across genres. Moreover, this study employs only two criteria: the degree of formality based on genre distributions and collocational patterns related to semantic preferences. To effectively differentiate near-synonyms, it is necessary to consider additional criteria, such as grammatical patterns (Phoocharoensil, 2010). It also points to exploring alternative statistical measures that are not limited to frequency and MI scores, such as the z-score, t-score, and log-likelihood tests. These measures could prove beneficial in analyzing the data in future research, considering the potential variations in noun and adverb collocates (Phoocharoensil, 2020). Overall, while this study provides valuable insights, it is necessary to recognize these limitations and address them in future studies in order to gain a more complete understanding of near-synonyms and their usage.

About the Author

Wimonnit Chaokongjakra: A lecturer at the Language Institute, Thammasat University, Thailand. She received her M.A. in Applied Linguistics (Teaching English for Specific Purposes) from Mahidol University. Her research interests include teacher professional development, students' and teachers' voices, EFL teacher education, and corpus linguistics.

References

- Alanazi, Z. (2023). A critical corpus-based analysis of the words Muslim and Islamic Vs. Christian in contemporary American English. *World Journal of English Language*, 13(2), 200-213.
- Aroonmanakun, V. 2015. Quick or fast: A corpus-based study of English synonyms. *Language Education and Acquisition Research Network (LEARN) Journal*, 8(1), 53–62. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/102685>.
- 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]. <http://www.diva-portal.se/smash/get/diva2:1017196/FULLTEXT01.pdf>
- Cai, J. (2012). *Is it "great" enough? A corpus-based study of "great" and its near synonyms* [Master Thesis]. Ball State University
- Carter, R. (2012). *Vocabulary. Applied linguistic perspective*. Routledge.
- Carter, R. & McCarthy, M. (1988). *Vocabulary and language teaching*. Routledge.
- Cheng, W. (2012). *Exploring corpus linguistics. Language in action*. Routledge.

- Chung, S. (2011). A corpus-based analysis of “create” and “produce”. *Chang Gung Journal of Humanities and Social Sciences*, 4(2), 399-425.
- Crystal, D. (2007). *Words words words*. Oxford: Oxford University Press.
- Davies, M. (2020). *The COCA corpus (new version released March 2020)*.
https://www.englishcorpora.org/coca/help/coca2020_overview.pdf
- Dvorak, J., Dawson, M. (2011). Crisis Management Terminological Synonyms. *Safety Engineering Series*, 6(2), 52-55.
- Edmonds, P., & Hirst, G. (2002). Near-synonymy and lexical choice. *Computational Linguistics*, 28(2), 105-144. <https://dx.doi.org/10.1162/089120102760173625>
- Folse, K. S. (2004). The underestimated importance of vocabulary in the foreign language classroom. *CLEAR News*, 8(2), 1-6.
<http://www.seasite.niu.edu/trans/articles/Underestimated%20Importance%20of%20Vocab.htm>
- Gablasova, D, Brezina, V. & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*. 67(1), 155-179.
- Gardner, D. (2013). *Exploring vocabulary in action*. Routledge.
- Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rded.). Routledge.
- Harley, H. (2006). *English words: A linguistic introduction*. Blackwell
- Imsa-ard, P. (2021). A corpus-based study on the meanings, distribution, collocations, and formality of ‘concentrate’, ‘emphasize’, and ‘focus’. *NIDA Journal of Language and Communication*, 26, 22-40.
- Jarunwaraphan, B., & Mallikamas, P. (2020). A corpus-based study of English synonyms: Chance and opportunity. *REFlections*, 27(2), 218–245. <https://so05.tci-thaijo.org/index.php/reflections/article/view/248710>
- Laufer, B. (1997). What's in a word that makes it hard or easy? Intralexical factors affecting the difficulty of vocabulary acquisition. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 140-155). Cambridge University Press.
- Lertcharoenwanich, P. (2023). Analysis of collocations and semantic preference of the near-synonyms: Blank, empty, and vacant. *LEARN Journal: Language Education and Acquisition Research Network*, 16(1), 365-383. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/263447>
- Longman Dictionary of Contemporary English. (Online).
<https://www.ldoceonline.com>
- Ly, T. H. & Jung, C. K. (2015). A corpus investigation: The similarities and differences of ‘cute’, ‘pretty’ and ‘beautiful’. *3L: The Southeast Asian Journal of English Language Studies*. 21(3), 125-140. <http://ejournals.ukm.my/3l/article/view/8788/3426>
- Martynska, M. (2004). Do English language learners know collocations? *Investigationes Linguisticae*, 11, 1-12.
- Merriam-Webster dictionary. (Online). <https://www.merriam-webster.com/>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
- Nisani, F. 2015. *A corpus-based study of English synonyms: Possible, probable, and likely*. [Master’s thesis, Thammasat University].
http://ethesisarchive.library.tu.ac.th/thesis/2015/TU_2015_5721040417_4760_2844.pdf
- Oxford advanced American dictionary. (Online).
https://www.oxfordlearnersdictionaries.com/definition/american_english/
- 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. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/111700>

- 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: Language, Linguistics, Literature. The Southeast Asian Journal of English Language Studies*, 26(3), 1–16.
- Phoocharoensil, S. (2020a). Collocational patterns of near synonyms error, fault and mistake. *The International Journal of Communication and Linguistic Studies*, 19(1), 1-17.
<https://dx.doi.org/10.18848/2327-7882/CGP/v19i01>
- Phoocharoensil, S. (2020b). A genre and collocational analysis of *consequence*, *result*, and *outcome*. *The Southeast Asian Journal of English Language Studies*, 26(3), 1 – 16.
<http://dx.doi.org/10.17576/3L-2020-2603-01>
- Phoocharoensil, S. (2022). Primary, main, and major: Learning the synonyms through corpus data. *GEMA Online Journal of Language Studies*, 22(4). 76–89.
<https://ejournal.ukm.my/gema/article/view/52706/13509>
- Saito, K. (2020). Multi- or single-word units? The role of collocation use in comprehensible and contextually appropriate second language speech. *Language Learning*, 70(2), 548–588.
- Sinclair, J. (2004). *Trust the text*. Routledge.
- 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 Publishers.
- Smyth, C. (2016). An introduction to corpus Linguistics. *Journal of Art and Sciences*, 14, 105–109.
<https://core.ac.uk/download/pdf/234563686.pdf>.
- Taylor, J. R. (2002). Near synonyms as co-extensive categories: ‘high’ and ‘tall’ revisited. *Language Science*, 25, 263-284.
- Thornbury, S. (2002). *How to teach vocabulary*. Longman.
- Wilkins, David. 1972. *Linguistics in Language Teaching*. Edward Arnold.
- Yang, B. (2016). A corpus-based comparative study of “learn” and “acquire”. *English Language Teaching*, 9(1), 209-220. <http://dx.doi.org/10.5539/elt.v9n1p209>