A Contrastive Corpus-based Study of American and English Adjectives: *'Important' and 'Significant'*

Bunyawat Sriwangrach

Kamphaeng Phet Rajabhat University, Thailand

Abstract

This contrastive corpus-based study aims to analyze the similarities and differences of two synonyms "important" and "significant" concerning on the degree of formality in their distribution across genres as well as their collocations and semantic preference. The corpus data derived from the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC) which represent varieties of English usage across different cultures (American and British). This study analyzed the top 20 noun collocates with Mutual Information (MI) scores value of at least three. The corpus data from COCA and BNC explored some shared characteristics and contrastive features. The findings discovered that the two adjectives have the same core meaning and degree of formality since their highest frequency in academic texts whereas frequencies are the lowest in informal genres i.e. TV and movie subtitles and fiction. In certain conditions, these synonyms are different in terms of noun collocation due to the fact that these synonyms co-occur with particular noun collocates. Interestingly, the findings also confirm two adjectives are near synonyms as they share only two similar theme of semantic preference. In consequence, the two synonyms cannot be used interchangeably in all contexts. Moreover, it was found that the corpora provide some useful details that are deficient in the dictionary.

Keywords: COCA, BNC, Genre Analysis, Noun Collocates, Semantic Preference

Introduction

English is extensively utilized on a global scale by people of different nations and has evolved into a common language for international communication (Harmer, 2007). Knowing English can lead students to achieve their goal and open career opportunities for those who use it to communicate effectively. In the process of acquiring proficiency in English, among four main skills, vocabulary is regarded as the paramount aspect of language learning as David Wilkins (1972) stated that "without grammar very little can be conveyed, without vocabulary nothing can be conveyed". This suggests that vocabulary is crucial for effective communication, as communication would be unattainable without a sufficient vocabulary. Nevertheless, to master English vocabulary is challenging due to the presence of numerous synonyms within the English language. Synonym is one of the difficulties that EFL leaners always encounter in vocabulary learning (Laufer, 1990). The EFL learners find it difficult to distinguish English synonyms clearly because of their similarities in meaning, i.e. the lexical relationship in which two or more linguistic forms have same meanings (Carter, 2012), British and American English synonyms could yield some difficulties to learners to distinguish the differences of the synonyms. In addition, limitations of the dictionary information could lead the difficulties in making a clear distinction among synonyms usage. To extend EFL vocabulary knowledge, it is noteworthy to raise awareness of synonym use. It is important to promote awareness of the limitations in substituting certain words for others, as not all words can be interchangeably used in every context.

Various factors and contexts need to be taken into account. Students tend to obtain more benefit in their learning progression if they possess an awareness of synonyms. Additionally, being aware of synonyms will improve English skills, especially writing abilities (Yeh et al., 2007).

Due to the fact that a dictionary cannot provide the accurate use of synonyms, the main focus of this present study is to investigate two synonym adjectives important and significant, with respect to their distribution across genre and collocations. In regard to the two synonyms mentioned earlier, the definitions and the examples of the use of those are extracted partially from two well-known and trusted online dictionaries, namely Oxford Advanced American Dictionary Online and Longman Dictionary of Contemporary English Online, were consulted for the purpose of comparison with the dictionary use.

Literature Review
Definitions of Synonym

A synonym, as described by Webb and Nation (2017), is a term or phrase that shares a similar meaning with another word. Various dictionaries offer additional explanations of the term. Oxford Advanced American Dictionary online defines a synonym as a word or expression with an identical or nearly identical meaning to another in the same language. The online Longman Dictionary of Contemporary English defines a synonym as a word that has the same meaning as another word in the same language.

Types of Synonyms

Synonyms can be classified into two main types: strict synonyms (absolute) and loose synonyms (relative). Synonyms are categorized into two types based on their interchangeability and contextual usage. Strict synonyms, as defined by Jackson and Amvela (2000), are words that can be substituted for each other across all contexts without altering the meaning, style, or connotation of the text. However, Kreidler (1998) argued that no two words possess entirely identical linguistic features, rendering complete interchangeability unnecessary as it results in redundancy. On the contrary, loose synonyms, as

described by Jackson and Amvela (2000), are words that overlap in meaning but cannot interchangeably replace in all contexts. Their interchangeability is limited to specific situations where their meanings align, without complete substitution across all instances.

To sum up, in comparison, loose-synonyms are possibly more common than strict synonyms. The incorrect use of synonyms can have an influence on the language's meaning, style, and context. Consequently, it is undeniable for the language learners to be able to distinguish and utilize synonyms appropriately.

Criteria for Distinguishing Synonyms

Jackson and Amvela (2007) have distinguished criteria to differentiate near-synonyms, e.g. dialect, degree of formality, collocation and semantic preference.

Dialect

Some synonyms are different in terms of meaning which can be distinguished by their various dialects. The Longman Dictionary of Contemporary English defines a dialect as a form of a language which is spoken only in one area. A dialect may be one of the national standard languages. For instance, 'lawyer' and 'attorney',British and American English synonym pair, 'lawyer' is spoken in British English, while 'attorney' is spoken in American English (Jackson & Amvela, 2000).

Degree of Formality

Another factor that can be used to distinguish between synonyms is degree of formality. In other words, a synonym is more suitable for utilization in formal contexts while the other is more commonly used in informal contexts. Jirananthiporn (2018) identified that the noun 'problem' are used in formal context, while 'trouble' seems to be more widely used in less formal context. In a similar vein, Phoocharoensil (2021) discovered that there seems to be a noticeable difference in formality between the verb synonyms "predict" and "foresee".With 'predict' occurring with the highest frequency in academic texts, on the contrary, "foresee" is widely used in webpage which is considered a less formal genre.

Collocation

Another critically vital criterion to distinguish near-synonyms is collocations. According to Baker et al. (2006) the word 'collocation' in corpus linguistics is defined as 'the phenomenon surrounding the fact that certain words are more likely to occur in combination with other words in certain contexts' (p.36). Furthermore, Flowerdew (2012) and Saito (2020) suggested that collocations are normally measured by statistical means, mutual information (or MI scores) etc. In other word, apart from looking the frequency, an MI score of 3 or higher should be taken into consideration that the two items are strongly collocated. By way of illustration, the synonyms cure and heal were investigated by Sumintana and Tawilapakul (2019). The corpus data from COCA were consulted and it revealed that cure is normally followed by disease like cancer, on the contrary, heal frequently seems to co-occur with something abstract such as wound or people.

Semantic Preference

In terms of semantic preference, Cheng (2012) pointed out that semantic preference is defined as the possibility for words to be limited to particular semantic circumstances. Flowerdew (2012) indicated that although collocation patterns are related with lexical constraint, the words' occurrences are determined by their semantic contexts. In other words, semantic preference can be seen as a semantic setting where vocabulary frequently occur. According to Edmonds and Hirst (2002), the verbs "die" and "pass away", although are equivalent, have distinct meaning preferences. The word "pass away" is only employed in connection to people; it is not used to refer to animals or plants.

As a result, collocations and semantic preference are closely related. Collocation patterns in a particular semantic context should be considered when deciding how to employ lexical elements appropriately. Consequently, this study analyzes linguistic features of near-synonyms important and significant, including distribution across genres, word frequency, typical noun collocates, and semantic preference by using corpus-based data from the Corpus of Contemporary American English (COCA) and British National Corpus (BNC) to identify their differences.

Corpus-based Studies on Synonyms

Several researches have conducted corpus-based studies on English synonyms to identify similarities and, more crucially, differences between them. It is evident that information derived from language corpora can offer fresh perspectives not found in learner dictionaries (Aroonmanakun, 2015). Rather than depending just on the dictionaries of native speakers, corpus-informed evidence can distinguish the distinctions between near-synonyms.

Chanchotphattha (2013) carried out a corpus-based study on three synonyms: different, diverse, and various,using data from the Oxford Advanced Learner's Dictionary (8th Edition International Student's Edition) (OALD) and concordance lines in the British National Corpus (BNC). The study's findings showed that the two regions of Britain and North America had distinct dialects that are frequently utilized with distinctive grammatical patterns. The research revealed that, in terms of formality levels, the context and particular meanings of these words determine whether they are formal or not. These adjectives cannot be replaced interchangeably in all contexts.

Another insightful study by Krawczak (2014) investigated the meanings of four English synonyms for "shame": embarrassed, humiliated, and ashamed. He conducted data analysis using a quantitative, corpus-driven methodology. British National Corpus (BNC) and the Corpus of Contemporary American English (COCA) were consulted. The findings demonstrated that these synonyms shared the same core meaning, on the other hand, there are differences in terms of collocation and sense of meaning.

Collocation is one of the main criteria used in differentiating the adjective synonyms. The synonym adjectives quick and fast were examine by Aroonmanakul (2015) to explore similarities and differences based on the information from COCA. The top 100 collocates of the chosen synonyms with MI scores of at least 3 were extracted. The results revealed that the two synonyms have different noun collocations. Quick commonly collocates with nouns that denote actions completed quickly, e.g. answer, breakfast, comment, reaction, reference, solution, trip, or visit. In contrast, fast typically happens when words that seem to describe the way something

moves, e.g. attack, acceleration, draw, boat, car, and etc.

In a similar vein, Petcharat and Phoocharoensil (2017) also confirmed the benefits of corpora over learner dictionaries in their study of synonymous adjectives appropriate, proper, and suitable, concentrating on meanings, degrees of formality, collocations, and grammatical patterns. The researchers examined these three synonyms by combining data from the Corpus of Contemporary American English (COCA) with entries from three learner dictionaries. It was found that although the three synonyms have the same core meaning, they, nonetheless, have different specific meanings, formality levels, collocations, and grammatical patterns. It was discovered that the corpus offers some insightful information not included in the three dictionaries.

From previous research, it indicates that employing corpus-based approaches for the analysis of synonymous terms can yield more thorough insights into colligation, collocation, distribution patterns, formality, and semantic preferences. Nevertheless, it seems there is a scantiness of corpus-based investigations into the distinctions and similarities between the adjectives significant and important. This research sought to examine the two adjectives in addressing the following research questions:

1. How are the synonyms important and significant distributed across different genres?
2. What are the common collocations of the adjective synonyms important and significant?

Methodology

In this contrastive corpus-based analysis, the near synonyms adjectives important and significant were examined. The online version of two dictionaries: Oxford Advanced American Dictionary and Longman Dictionary of Contemporary English provided the fundamental definitions of the target synonyms. The two target synonyms were analyzed using two significant corpus-based sources. First, the Corpus of Contemporary American English (COCA) was consulted. COCA is one of the most well-known and extensively used in the field of ELT for research and practice representing American English. As a

mega-balanced corpus, COCA contains more than one billion words of texts, with around 25 million words added each year from 1990-2020. COCA is evenly divided in eight various genres namely, TV/movie subtitles, spoken, fiction, popular magazines, newspapers, blogs, web pages, and academic texts, which cover both spoken and written texts. The second corpora, British National Corpora (BNC) represents British English. To expound, BNC was originally created by Oxford University Press in the 1980s and early 1990s containing 100 million words of texts from wide range genres, i.e., spoken, fiction, magazines, newspapers and academic texts. The information gathered from COCA and BNC was triangulated in order to look at the differences between the two near synonyms in various English dialects. An extensive and in-depth analysis was conducted on the target synonym distribution across genres in COCA and BNC. To answer the two above research questions,frequencies and distribution patterns of the target synonyms, i.e. important and significant, in all of the eight different genres of COCA and five different genres of BNC, were first explored in order to answer the first research question. As for the second question, the majority of noun collocates were focused on because the target synonyms are adjectives, therefore, the noun collocates frequently accompanying the two nouns were investigated. The frequency and mutual information (MI) score were taken into consideration when choosing the noun collocates. In addition, since the co-occurrence of terms is statistically significant, the corpus statistics were utilized to determine the strength of collocations, with a minimum MI score of 3 (Hunston, 2002). Furthermore, Schmitt (2010) recommended utilizing the MI score with caution along with the frequencies. The first top twenty noun collocates of the target synonyms from COCA and BNC which frequently co-occur with the highest frequencies and the significant MI score level ($\geq$ 3) were taken. Following that, the target collocates were scrutinized according to their semantic preference. In other words, the similar collocations were grouped under a common theme. Finally, the sentences or phrases from the concordance lines were extracted to exemplify how each synonym co-occurs in particular contexts.

Results and Discussion

Based on the criteria employed in this study, the results of the investigation are examined in terms of definitions, degree of formality, and collocation.

Table 1 The Definitions and Examples of "Important" and "Significant" derived from Online American – English Dictionary

Synonyms	Oxford Advanced American Dictionary	Longman Dictionary of Contemporary English
1. important	**Definition** having a great effect on people or things; of great value E.g. I have an *important* announcement to make. E.g. Money played an *important* role in his life.	**Definition** An important event, decision, problem etc. has a big effect or influence on people's lives or on events in the future E.g. The accident taught him an *important* lesson. E.g. It's *important* to explain the procedure to the patient.
2. significant	**Definition** large or important enough to have an effect or to be noticed E.g. There are no *significant* differences between the two groups of students. E.g. Your work has shown a *significant* improvement	**Definition** having an important effect or influence, especially on what will happen in the future E.g. Please inform us if there are any *significant* changes in your plans. E.g. The rise in temperature is not statistically *significant*.

The Table 1 illustrates the definitions that important and significant are similar in meaning and they can be considered as near-synonyms of one another. Nonetheless, the selected synonyms share overlapping notions. The information in terms of their use in various genres, the degree of formality and collocation are not explained. As a result, significant and important seem to be problematic dictated the choice of synonyms among EFL learners due to the lack of a clear distinction among these near-synonyms. Therefore, this corpus-based analysis aims at investigating the similarities and differences of these two target synonyms, moreover, insightful information on specific patterns and usages in natural settings are provided so that the EFL learners can utilized them in appropriate contexts (Lee & Liu, 2009; Ly & Jung, 2015).

In response the two research questions, after presenting the data on the general frequency of the two target synonyms in eight distinct genres from COCA and five different genres from BNC, the collocations with the synonyms that frequently occur are shown.

Table 2 Distribution of Synonyms Important and Significant across Eight Genres According to their Frequency of Occurrence derived from COCA

Important			Significant		
Genre	Frequency	Per million	Genre	Frequency	Per million
academic texts	76,605	639.49	academic texts	60,470	504.80
spoken	51,467	408.03	webpages	15,190	122.25
blogs	46,809	363.95	blogs	13,194	102.59
webpages	46,725	376.05	magazines	11,590	91.92
magazines	41,134	326.22	newspapers	10,700	87.89
newspapers	29,000	238.21	spoken	9,002	71.73
TV and movie subtitles	25,290	197.46	fiction	1,711	14.46
fiction	14,649	123.81	TV and movie subtitles	1,247	9.74
Total	331,679		Total	123,104	

The Table 2 clearly illustrates the total frequency of the two adjective synonyms in COCA. The highest frequency derived from the adjective important (331,679 tokens) is over two times higher than that of significant (123,104 tokens). It revealed that both adjectives important and significant are associated in some characteristics in which they are common in academic texts (76,605 tokens and 60,470 tokens) and this yield a very high degree of formality. This denotes that the usage of important and significant in American formal contexts is common. As can be seen in the top 5 genre, most of which represent in written genre. On the contrary, the second rank of both adjectives seems to be appeared in different aspects. To manifest, the adjective important occurred in spoken genre to be the second rank (51,467) whereas, significant occurred in webpages within the written genre (15,190 tokens). Apart from the second rank, these two synonymous adjectives have shared some characteristics in which both spoken and webpages are considered as informal and colloquial language which represents less level of formality. In particular, the lowest frequency of important is in fictions (14,649 tokens) as well as significant appears to be the lowest frequency in TV and movie subtitles (1,247 tokens). Based on their low frequency in TV and movie subtitles and fictions which is considered as informal language, it can be stated that they are not commonly associated with colloquial language.

The following section demonstrates the analysis of the common noun collocates of the synonymous adjectives important and significant from COCA listed in Table 3 in responding to the second research questions. The total frequency of the target synonyms is shown based on the genres in which they appear.

Table 3 Noun Collocations of the Synonyms Important and Significant in COCA Across Eight Genres

Rank	Important			Significant		
	Noun collocates	Frequency	MI score	Noun collocates	Frequency	MI score
1	role	6,671	3.76	difference	11,513	5.97
2	factor	4,339	3.71	effect	4,632	4.42
3	aspect	2,401	3.81	change	3,626	3.74
4	lesson	1,487	3.07	amount	2,456	4.27
5	component	1,315	3.39	impact	2,104	4.33
6	implication	1,101	3.74	role	2,009	3.46
7	consideration	1,087	3.68	increase	1,786	4.66
8	contribution	1,034	3.15	relationship	1,756	3.21
9	distinction	902	3.88	correlation	1,740	6.32
10	determinant	316	5.17	improvement	1,559	5.20
11	predictor	267	3.53	interaction	1,498	5.11
12	milestone	261	4.30	factor	1,483	3.59
13	caveat	150	4.03	contribution	1,326	4.93
14	constituency	126	3.01	portion	1,247	5.12
15	facet	114	3.31	predictor	1,241	7.18
16	takeaway	61	3.41	reduction	1,057	5.01
17	pollinator	32	3.75	variable	996	4.47
18	correlate	31	3.25	gender	869	4.20
19	pathogenesis	31	3.95	association	798	3.10
20	Personage	18	3.34	progress	709	3.67

The Table 3 presents the nouns which commonly collocate with the adjectives important and significant as determined MI value of ($\geq$ 3) at the minimum, i.e. the level of collocational association's statical

significance, are listed in Table 3. As near-synonyms whose core meaning are closed, obviously, there are some noun collocates shared by the two adjectives. It can be stated that the selected adjectives are synonymous of each other (Phoocharoensil 2020a, 2020b). As shown in Tale 3, the data from the corpus indicate the near synonymous status of important and significant as there are some of nouns are frequently co-occurred with such as role, factor and predictor. To manifest, with a limitation of this study, only the top-20 nouns were targeted in the extraction of collocations. Though their frequency or MI score in COCA is less than 3, it appears that there are several nouns that can co-occur with the two target synonyms but are not included in either of the two lists. For instance, the number of occurrences of important finding is high in COCA (869 tokens), but its MI score is 2.45 which is lower than 3 resulted in its exclusion from the collocation list.

In the following section, the list of the noun collocation as shown in Table 3 provided the noun collocates of the synonymous adjectives important and significant will be subsequently categorized based on their semantic preferences.

Table 4 Semantic Preference of Noun Collocates of Important from COCA

Semantics preference	Noun collocates
1. Prediction	predictor, caveat
2. Advancement	milestone, lesson
3. Duty	role, constituency, personage
4. Connection	correlate, pollinator
5. Divergence	distinction, aspect
6. Cause	determinant, factor, consideration
7. Research-Related Matters	contribution, implication, facet, component
8. Miscellaneous	pathogenesis, takeaway

According to the analysis of semantic preference of noun collocates of important from Coca in Table 4, eight principle themes were discovered. The first theme Prediction includes predictor and caveat, as exemplified in (1), and the second theme, Advancement, includes milestone and lesson, as shown in (2). Third theme refers to Duty including role, constituency and personage, as exemplified in (3). Connection, the forth theme, containing

correlate and pollinator, as shown exemplified in (4). The fifth theme consists of distinction and aspect, as exemplified in (5), is Divergence. The sixth theme contains three nouns namely, determinant, factor and consideration is CAUSE as in (6). The seventh theme, containing contribution, implication, facet and component, as in (7), is Research-Related Matters. The two nouns that do not seem to fit in any themes, i.e. pathogenesis and takeaway, have been assigned to Miscellaneous.

1. Static acceleration was one of the most important predictors of behavior in the favored random forest model classifying our 6 behavioral states.
2. It was signed at a conference on 14 April 1961 and is an important milestone in the history of international relations.
3. They're looking for strong leadership and I think that's a very important constituency right now in America.
4. Research suggests that unwanted social exclusion is an important correlate of anxiety and depression
5. Diplomatic is important aspect. It's a military aspect.
6. In our study, individual factors were the most important factors related to having a dental check-up in the previous year for Iowa children.
7. Our findings highlight the important contribution that leisure-time physical activity in adulthood can make to longevity.

Table 5 Semantic Preference of Noun Collocates of Significant from COCA

Semantics preference	Noun collocates
1. Prediction	predictor
2. Fluctuation	difference, increase, reduction, change
3. Duty	role
4. Connection	relationship, correlation, interaction, association
5. Measurement	amount, portion
6. Effect	impact, effect
7. Research-Related Matters	improvement, contribution, variable, progress, factor
8. Miscellaneous	gender

Through an investigation into the semantic preference of noun collocation of significant, eight major themes are presented. The first theme, Prediction contains only one finding, predictor, as shown in (8), the second theme, Fluctuation, contains noun collocation that are similar in sense, i.e. difference, increase, reduction and change, as exemplified in (9). The third theme, Duty, contains only one noun collocation, role, as in (10) and Connection, including relationship, correlation, interaction and association, as in (11). The fifth theme, containing amount and portion, as exemplified in (12), is Measurement. Effect comprised of impact and effect, as shown in (13). The majority of the noun collocation, the seventh theme, fell into Research-Related Matters which consists of improvement, contribution, variable, progress and factor, as exemplified in (14). The noun, gender, does not seem to fit in any theme, therefore, has been assigned to Miscellaneous.

1. High school GPA, SAT scores, and college major are significant predictors of final college.

2. Significant increases in hemoglobin in anemic children have also been reported after daily dosing with vitamin.
3. He feels that the federal government should have a significant role in day to day life.
4. No statistically significant relationships were found in the total or subset responses.
5. I bet he would increase payroll a significant amount.
6. Multiple comparisons were conducted to determine the significant effects by the least significant difference.
7. The results showed that there was a statistically significant improvement in glycosylated hemoglobin.

On the contrary, the distributions across five different genres in British National Corpus (BNC), namely academic, newspaper, magazine, fiction and spoken were analyzed to represent British English. As a normalized frequency, the results were ranked according to frequency per million.

Table 6 Distribution of Synonyms Important and Significant According to Frequency BNC across Five Genres

Important			Significant		
Genre	Frequency	Per million	Genre	Frequency	Per million
Academic	10,708	698.42	Academic	4,813	313.93
magazine	2,225	306.39	Newspaper	662	63.25
spoken	2,886	289.65	magazine	358	49.30
newspaper	2,169	207.23	spoken	403	40.45
fiction	2,307	145.01	fiction	183	11.50
Total	20,295		Total	6,419	

It can be seen from the Table 6 that the total frequency of the two adjective synonyms, overall, important occurs with frequency in BNC (20,295 tokens) which higher than significant (6,419 tokens). Interestingly, this sequence is aligned with the total frequency in COCA. According to Table 6, important is found the most frequency in academic due to the high rate of occurrence (10,708 tokens), followed by magazine (2,225 tokens). While, significant is found the most frequency in academic (4,813 tokens) followed by newspaper (662 tokens).It is clear that these two synonyms are common in written context as can be seen in top two genres. This indicates that the two adjectives synonyms are common in formal English. On the contrary, based in the evidence, it is apparent that the adjectives important and significant have the lowest frequency in fiction (2,307 tokens and 183 tokens) which associated with informal or colloquial language. In similar vein, the results align with the occurrence of important and significant in American fiction contexts which have the lowest frequency. Therefore,the two adjectives occur with the greatest frequency in academic texts among the other genres. Their high frequencies of the adjectives important and significant in academic contexts can support the assumption that the two synonyms have a high level of formality. It can be concluded that the usage of the synonymous adjectives important and

significant in British formal contexts is common.

As a result, it can be summarized that the two synonymous adjectives in both COCA and BNC have a high degree of formality. The word register or style of the settings in which these synonyms are used varies. Furthermore, due to the similarity in the analysis, the status of these synonyms adjectives can be confirmed as near-synonyms.

The following section will focus on the collocational differences and similarities of the synonyms important and significant. Similar to the analysis from COCA, the MI score higher than 3 also will be taken into account as the criteria to show that the target synonyms are significantly collocates in BNC.

Table 7 Comparison of Frequency of the Synonyms Important and Significant in Different Text Type in BNC across Five Genres

Rank	Important			Significant		
	Noun collocates	Frequency	MI score	Noun collocates	Frequency	MI score
1	role	761	3.74	difference	419	5.26
2	factor	533	4.76	difference	366	5.59
3	issues	339	3.14	changes	245	3.72
4	factors	294	3.42	increase	197	3.58
5	aspect	293	4.42	proportion	182	4.82
6	feature	252	3.68	contribution	176	5.08
7	element	227	3.70	role	150	3.09
8	contribution	183	3.44	impact	142	4.29
9	implications	177	3.64	factor	139	4.51
10	consideration	159	3.21	correlation	117	6.43
11	distinction	122	3.23	improvement	98	4.59
12	component	98	3.56	numbers	97	3.15
13	respects	84	4.02	reduction	84	4.16
14	considerations	64	3.10	progress	79	3.33
15	ingredient	40	4.44	developments	64	3.68
16	determinant	32	5.22	improvements	57	4.59
17	insights	28	3.53	feature	57	3.23
18	indicator	28	3.42	increases	47	3.51
19	discoveries	24	3.50	minority	43	3.70
20	distinctions	23	3.30	reductions	42	5.04

The Table 7 represents the top-20 of several nouns strongly collocates with the adjectives important and significant as determined by the MI score of (≥ 3) found in BNC. It is apparent from the Table 7 that despite the adjectives important and significant are considered as near-synonyms due to the fact that their core meanings are closed, only a small number of noun collocations are overlapping shared by these adjectives, namely, role, factor (s), feature and contribution. Obviously, there are more varieties in the collocations derived from BNC than from COCA. It is clear that the first most frequent

noun collocations of the adjectives important and significant, i.e. role and difference, are similar to the most frequent nouns that discovered in COCA. Apparently, the noun role is the most frequent noun collocates of important and different is the most frequent noun collocates of significant.

Nonetheless, there are many noun collocates outside these source presented in the present study and are excluded from the collocation lists due to the fact that those nouns have the MI score lower than 3 which do not align with the criteria used in the selection of the study.

The next section presents the noun collocates of the synonymous adjectives important and significant those were subsequently distributed based on their semantic preference.

Table 8 Semantic Preference of Noun Collocates of Important from BNC

Semantics preference	Noun collocates
1. Composition	element, component, ingredient
2. Cause	factor, factors, determinant, issue
3. Divergence	distinction, distinctions, aspect
4. Thought	consideration, considerations, insight, respect
5. Duty	role
6. Research-Related Matters	contribution, implications, indicator, discoveries
7. Miscellaneous	feature
8. Miscellaneous	gender

Upon investigation of the semantic preference of the noun collocation of adjective important, seven principle themes were determined. As revealed in Table 8, the first theme Composition includes element, component and ingredient as shown in (14) followed by the second theme Cause which contains factor, factors, determinant and issue, as exemplified in (15). The third theme Divengence has three members, namely distinction, distinctions and aspect, as shown in (16). The forth theme containing four nouns collocation i.e. consideration, considerations, insight and respect, as exemplified in (17), is Thought. The next theme has only one key noun collocation, role, is Duty (18). There are exist certain four noun collocations i.e. contribution, implications, indicator, and discoveries which is related to Research-Related Matters, as exemplified in (19). The last noun, feature, which does not seem to belong to any themes, is placed in Miscellaneous.

1. Liquidity is normally an important element of RIE status.
2. One of the most important factors is the interest rate itself, which could give cause for confusion.
3. This subculture had its own important distinctions.
4. These are important considerations to keep in mind when designing Compact goals.
5. God plays an important role in Coleridge's

poetry, as a symbol of ultimate creativity.
6. They make an important contribution to the intake of vitamin C and potassium.

Table 9 Semantic Preference of Noun Collocates of Significant from BNC

Semantics preference	Noun collocates
1. Fluctuation	increase, increases, reductions, change, different, differences, reduction
2. Advancement	developments, progress
3. Duty	role
4. Connection	correlation
5. Measurement	proportion, numbers
6. Subgroup	minority
7. Research-Related Matters	contribution, improvements, improvement, impact, factor
8. Miscellaneous	feature

As demonstrated in Table 9, eight themes of the noun collocates of important from BNC have been assigned. The first theme, Fluctuation, the majority of noun collocates, contains increase, increases, reductions, change, different, differences and reduction as seen in the example (14). Second theme, Advancement which includes two nouns collocate e.g. developments and progress, as exemplified in (15). The third theme related to important and significant from Coca is Duty, containing only one key noun role, as exemplified in (16). Connection, the forth theme contains correlate, as in (17). The fifth theme is relevant to Measurement, as exemplified in (18). Next, sixth theme, Subgroup, contains one collocation, minority, as in (19). The seventh theme, containing contribution, improvements, improvement, impact and factor, as exemplified in (20), fell into Research-Related Matters. One noun, feature, does not seem to fit in any theme, has been assigned to Miscellaneous. Notably, It is important to remember that further possible noun collocates may reveal new themes in future studies.

1. There must be a significant increase in pensions and benefits in order for them to be able to pay.
2. One of the most significant developments in the past decade has been the growth of women's alternative media.

3. They will retain a significant role through the party machinery in policy formation.

4. There was no significant correlation between arterial P co 2 and alcohol consumption.

5. Immigration from the West Indies began to reach significant numbers.

6. Disabled adults of working age are a significant minority group, experiencing a rate of unemployment.

7. A significant improvement was found with active treatment but not with placebo.

In summary, a thorough analysis revealed differences among the target synonyms. The finding of this study related with previous studies in many aspects. In terms of genres, the two adjectives synonyms important and significant are widespread in academic texts which substantiate their high level of formality. A closer investigation at the nouns that commonly co-occur with the adjectives important and significant from COCA and BNC seems to provide insightful information on the usage of distinction of both adjectives. Through a closer investigation into the semantic preference, it appears that the target synonyms do not share all the same noun collocates, noun collocates with the two target adjectives important and significant share only two themes in common, namely, Duty and Research-Related Matters, it can be concluded that the collocational patterns of important and significant from Coca and BNC can strengthen their status as near-synonymy because they are unable to share subtleties of meaning. Furthermore, the current findings of this study in line with Chanchotphattha (2013) who stated that the two regions of Britain and North America have distinct dialects that are frequently utilized with distinctive patterns. Additionally, the finding also in line with Petcharat and Phoocharoensil (2017), Selmistraitis (2020) and Song (2021) that synonymous words show some distinctions in terms of collocation and semantic preference even they share the same core meaning and degree of formality, resulting that they cannot be used interchangeably in all contexts (Cai, 2012; Hu, 2015; Phitayakorn, 2016; Phoocharoensil, 2010). It can be stated that the corpus offers some insightful information which is non-exist in the dictionaries.

Conclusion and Recommendations

In conclusion, this contrastive study sheds some light on how to implement corpus data to analyze English synonymous adjectives. The synonymous adjectives important and significant were investigated to demonstrate similarities and differences between these two words based on their degree of formality, collocational patterns with respect to semantic preferences. The data were drawn from two online dictionaries, Oxford Advanced American Dictionary and Longman Dictionary of Contemporary English, and were analyzed both quantitatively and qualitatively by using COCA and BNC which represent English usage in different contexts (American and British). The results have clearly demonstrated that both adjective synonyms share the same core meaning and the two synonyms happen to be favored in formal genres, their highest frequency represented in academic texts rather than spoken discourse. This analysis is further reinforced by the low frequency of occurrences in informal genres such as fiction and TV/movie subtitles. Even though these target synonyms have a similar degree of formality, these synonyms tend to have quite different occurrence of noun collocation.

This study also presents the top-20 of nouns which most strongly collocate with these synonyms from both COCA and BNC. Based on the noun collocation of the target synonyms, lexical patterns and usages in naturally occurring settings are shown by differences in noun collocates. Thus, there are several modest use variations among the synonyms, as shown by a comprehension analysis of the wide semantic preference of noun collocation. It can be concluded that the adjectives important and significant cannot be used interchangeably in all contexts. From the semantic preference analysis, it can be seen that several noun collocates from BNC which represents British English usage appeared to be in plural forms i.e., factors, distinctions, considerations, implications, increases, reductions, differences, developments, numbers and improvements, whereas, the American utilized the target synonyms with only singular noun collocation. It is trustworthy to note that the British and American language users utilized these synonyms differently in some perspectives. This interesting result confirms their status as near-

synonyms. For EFL students, this might be more advantageous than just consulting a dictionary or depending only on the intuition of native speakers which isn't always accurate and trustworthy. The corpus tools can shed more light a new research perspective to expose language use beyond the intuitive perception of a native speaker (Hunston, 2002).

Nonetheless, there are certain restrictions on the present study. Only two synonym adjectives are examined and the number of noun collocates is restricted to those found in the top 20 lists. In order to obtain more interesting results, it could be beneficial for future studies by looking at more near-synonyms utilizing various corpora. Moreover, distribution and collocation pattern were the primary focus of this study, grammatical patterns can also be taken into account (Phoocharoensil, 2010). The finding may not be generalized to include other varieties of English inner-circle Englishes, namely Australian English, or in the ESL contexts, e.g. Indian English, Chinese English, Singapore English etc.

References

Aroonmanakun, V. (2015). Quick or fast: A corpus based study of English synonyms. *LEARN Journal: Language Education and Acquisition Research Network, 8*(1), 53-62.

Baker, P., Hardie, A., & McEnery, T. (2006). *A Glossary of Corpus Linguistics*. Edinburgh University Press.

Cai, J. (2012). *Is it "Great" Enough? A Corpus-based Study of "Great" and its Near Synonyms*. Ball State University.

Carter, R. (2012). *Vocabulary: Applied Linguistic Perspectives*. Routledge.

Chanchotphattha, K. (2013). *A Corpus-Based Study of English Synonyms: Different, Diverse, and Various*. Thammasat University.

Cheng, W. (2012). *Exploring Corpus Linguistics: Language in Action*. Routledge.

Edmonds, P., & Hirst, G. (2002). Near-synonymy and lexical choice. *Computational Linguistics, 28*(2), 105-144.

Flowerdew, L. (2012). *Corpora and Language Education*. Springer.

Harmer, J. (2007). *The Practice of English Language Teaching*. Pearson Education.

Hu, H. M. (2015). A semantic prosody analysis of three adjective synonymous pairs in COCA. *Journal of Language and Linguistic Studies, 11*(2), 117-131.

Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge University Press.

Jackson, H., & Amvela, E. (2000). *Words, Meaning and Vocabulary*. A&C Black.

Jackson, H., & Amvela, E. (2007). *Words, Meaning, and Vocabulary: Introduction to Modern English Lexicology*. Continuum.

Jirananthiporn, S. (2018). Is this problem giving you trouble? A corpus-based examination of the differences between the nouns 'problem' and 'trouble'. *Thoughts,* (2), 1-25.

Krawczak, K. (2014). *Shame and its Near-Synonyms in English: A Multivariate Corpus-Driven Approach to Social Emotions*.

Kreidler, C. W. (1998). *Introducing English Semantics*. Routledge.

Laufer, B. (1990). Words you know: How they affect the words you learn. In J. Fisiak (Ed.), *Further Insights into Contrastive Linguistics (*pp. 573-593). John Benjamins.

Lee, C., & Liu, J. (2009). Effects of collocation information on learning lexical semantics for near synonym distinction. *Computational Linguistics and Chinese Language Processing, 14*(2), 205-220.

Liu, D., & Jiang, P. (2009). Using a corpus-based lexico grammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal, 93*(1), 61-78.

Ly, T. H., & Jung, C. K. (2015). A corpus investigation: The similarities and differences of cute, pretty and beautiful. *3L: Language, Linguistics, and Literature, 21*(3), 125-140.

Petcharat, N., & Phoocharoensil, S. (2017). A Corpus-Based Study of English Synonyms: Appropriate, Proper, and Suitable. *LEARN Journal: Language Education and Acquisition Research Network, 10*(2), 10-24.

Phitayakorn, W. (2016). *A Corpus-based Study of English Synonyms: Advise, Recommend, and Suggest*. Thammasat University.

Phoocharoensil, S. (2010). A corpus-based study of English synonyms. *International Journal of Arts and Sciences, 3*(10), 227-245.

Phoocharoensil, S. (2020a). Collocational patterns of the near-synonyms error, fault, and mistake. *International Journal of Communication and Linguistic Studies, 19*(1), 1-17.

Phoocharoensil, S. (2020b). A genre and collocational analysis of consequence, result, and outcome. *3L: The Southeast Asian Journal of English Language Studies, 26*(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.

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.

Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Springer.

Selmistraitis, L. (2020). Semantic preference, prosody and distribution of synonymous adjectives in COCA. *GEMA Online Journal of Language Studies, 20*(3), 1-18.

Song, Q. (2021). Effectiveness of corpus in distinguishing two near-synonymous verbs: Damage and Destroy. *English Language Teaching, 14*(7), 8-20.

Sumintana, N., & Tawilapakul, U. (2019). A corpus-based study of the English synonyms cure and heal. *8th LITU International Graduate Conference*.

Wilkins, D. A. (1972). *Linguistics in Languages Teaching*. Edward Arnold.

Webb, S., & Nation, P. (2017). *How Vocabulary is Learned*. Oxford University Press.

Yeh, Y., Liou, H-C., & Li, Y-H. (2007). Online synonym materials and concordancing for EFL college writing. *Computer Assisted Language Learning, 20*(2), 131-152.

Author Details

Bunyawat Sriwangrach, *Kamphaeng Phet Rajabhat University, Thailand,* **Email ID***: bunyawat_3j@hotmail.com*