

# A Corpus-Based Study Of English Synonyms Of The Adjectives ‘Far’, ‘Distant’, And ‘Remote’

Niramol Thongpan

*Language Institute of Thammasat University, Thailand, Niramol.thp@gmail.com*

## Abstract

This corpus-based study investigates the similarities and differences between the three synonymous adjectives ‘far’, ‘distant’, and ‘remote’, concentrating on the degree of formality in their distribution across eight genres, as well as their collocations with the semantic preference combination. The data for this study was gathered by using COCA (Corpus of Contemporary American English). This study also analyzed the top 30 noun and adverb collocates with Mutual Information (MI) scores of at least three. The results revealed that the word ‘remote’ has the most formal degree, followed by ‘distant’ and ‘far’, respectively, since it is most commonly used in academic texts. In terms of collocations and semantic preference, the data pointed out that the three synonyms share only a few collocates and, as a result, they vary in semantic preference. Interestingly, ‘far’ appears to be uncommon when used with an adverb, compared to their synonyms. Consequently, despite being near-synonyms, they cannot be perfectly substituted for each other in all contexts. More importantly, the results of the study will encourage English teachers to apply this information pertaining to the different frequency across genres and collocations integrated with semantic preference, when designing the pedagogical materials for students to learn how to distinguish between synonyms and correctly apply them in contexts.

**Keywords:** synonym, degree of formality, genre, collocation, semantic preference

## 1. Introduction

Vocabulary is considered the most important constituent in learning English (David, 1972, p.111). Without vocabulary knowledge, the communication would not be successful. Laufer (1990) indicated that vocabulary learners are normally challenged with learning synonyms. A synonym is defined as a word or expression which serves the same or almost the same meaning as another in the same language (Oxford Advanced Learner's Online Dictionary). The adjectives ‘proper’ and ‘suitable’ is an example of a synonymous pair. In addition, there are many times when students seem to make mistakes unintentionally when they use synonyms. To be more specific, they apply a synonym to replace another in all contexts, leading to ungrammaticality or less natural language (Szudarski, 2018; Thornbury, 2002). Although some of them may acknowledge that problem and they then depend on a dictionary to help, they are not

provided all the answers, e.g. their frequency of use in different genres, degree of formality and collocations, due to the limitation of a space in a book.

In this study, three adjective synonyms, namely, ‘far’, ‘distant’ and ‘remote’ were selected since they are very close in terms of meaning, as a result, users probably get confused and apply in contexts improperly. Moreover, the three words are even common vocabulary. The word ‘far’ is used with a high frequency and the two synonymous others are used at the medium level, as indicated by Longman Dictionary of Contemporary English Online. Furthermore, these words have been examined by only one researcher, as appeared in the study by Podzimek (2008). However, the study has not been updated to the present. The study only revealed the results of the analysis of noun collocations of the target words, other aspects were not examined.

In addition, the definitions and the examples of the use of the target words were derived from 2 well-known and trusted American-English

dictionaries, namely Oxford Advanced American Dictionary, and Merriam-Webster Dictionary, and demonstrated in Table 1 below.

**Table 1: Definitions and examples of ‘far’, ‘distant’ and ‘remote’ derived from 2 American-English dictionaries**

Synonyms	Oxford Advanced American Dictionary	Merriam-Webster dictionary
1. <b>far</b>	at a greater distance away from you e.g. I saw her on the <b>far</b> side of the road.	remote in space such as the <b>far</b> reaches of outer space
2. <b>distant</b>	far away in space or time such as <b>distant</b> stars/planets	separated in space: away such as a mile <b>distant</b>
3. <b>remote</b>	far away from places where other people live such as a <b>remote</b> beach	separated by an interval or space greater than usual such as an involucre <b>remote</b> from the flower

As can be observed from Table 1, ‘far’, ‘distant’ and ‘remote’ share the same meaning of being far away, especially in terms of space. However, the information in terms of their use in various genres, the degree of formality as well as collocations does not exist. As a result, this study aims to investigate the similarities and differences of these target synonyms regarding the formality degree based on the distribution across genres and their collocations integrated semantic preferences by employing the process of Corpus software, COCA.

## 2. Literature review

### 2.1 Definition of synonym

A synonym is defined as a word or phrase which has the same meaning as another one (Webb and Nation, 2017). Despite considered synonyms, they can differ in some contexts. Some examples of the synonym pairs are provided by Collins Thesaurus Online, such as ‘collect’ and ‘gather’, and ‘reject’ and ‘deny’, etc.

### 2.2 Types of Synonyms

Synonyms can be categorized into two kinds, namely ‘strict’ and ‘loose’ synonyms. A strict synonym is a kind of synonym that a word can be substituted for another without changing its

original meaning, style, or connotation in all possible contexts (Cruse, 1986). Due to this characteristic, it can definitely result in redundancy of language and, as a result of that, a word replaced by the other may finally turn out to disappear. ‘Fain’ is one example of obsolete words while its synonym ‘willing’ has still been used nowadays (Jackson & Amvela, 2000, p.94). However, it is fortunate since there are not many of that kind of synonym.

Another sort of synonym is a loose synonym which refers to a word that cannot be replaced in all contexts perfectly. Replacing them in all contexts results in significantly overlapped meaning. Although synonyms share a meaning, they can be different in terms of meaning in detail as well as other aspects. Jackson and Amvela (2000, p. 94) provide the examples that ‘discover’ and ‘find’ are synonyms whose meanings differ in detail. To illustrate this point, ‘discover’, as in ‘Marie Curie discovered radium in 1898.’ means ‘to find something at the first time’, on the other hand ‘found’, the past simple form of ‘find’, denotes ‘to experience something in some way’ as shown in ‘Franz found it easy to compose sonatas’. What is more, since this kind of synonym exists with a much larger amount than a strict synonym, it is then crucial for users who should

make understanding their difference to apply in contexts properly. In the next part, the criteria for helping distinguish synonyms are revealed.

### 2.3 Criteria for Distinguishing Synonyms

To understand synonyms which have some difference in some ways, five criteria for helping differentiating between synonyms are provided below.

#### 2.3.1 Dialect

Some synonyms are different in terms of dialect. To be more specific, a synonym is used in one area while the other is used in another. For instance, ‘splinter’ is used in standard English, while its pair ‘skelf’ is spoken in Northern British English. More interesting example regarding national standards is the pair synonyms ‘tap’ and ‘faucet’, the former is used in standard English, while the latter is used in American English (Jackson and Amvela, 2000, p.95).

#### 2.3.2 Degree of formality

Degree of formality is another criterion that can be used to differentiate between synonyms. That is, a synonym is more suitable to apply for formal contexts, in contrast, the other is better if used in informal contexts. Phoocharoensil (2010) identified that ‘request’ and ‘appeal’ are used in a more formal context than ‘ask’, ‘beg’, and ‘plead’ since ‘appeal’ tends to occur with the words expressing formality and ‘appeal’ involves the contexts of serious public request or a formal request to a court or someone in authority. More example provided by Jackson and Amvela (2000, p.95) is the pair synonym ‘drunk’ used as standard English, while ‘sloshed’ is considered a slang synonym.

#### 2.3.3 Connotation

The third way for distinguishing synonyms is to consider connotations. This method relates to synonyms that have different meanings in terms of emotion. To elaborate, one synonym expresses a positive feeling, while the other expresses a negative feeling. For instance, ‘famous’ is used to describe somewhere, something, or someone being well known in a positive context, as in ‘he became internationally famous for his novels’ from the Oxford Advanced American Dictionary,

whereas its synonyms ‘notorious,’ as in ‘a notorious computer hacker or notorious cases of human rights abuses’ from the Longman Dictionary of Contemporary English, online version, are used in negative context.

#### 2.3.4 Collocation and semantic preference

Collocation integrated with semantic preference is another way to help confirm the difference between synonyms, as indicated in the study of Szudarski (2018). Collocation refers to words that are often used together (Longman Dictionary of Contemporary English, 2014, p. 336). Moreover, the word meaning in collocation is determined not only by the word itself, but also by how it occurs with other words around them. For example, ‘pay’ usually appears with ‘attention’, while ‘commit’ frequently collocates with ‘crime’ (O’Dell and McCarthy, 2008).

More interesting example is the synonym pair ‘rancid’ and ‘sour’, despite sharing the meaning as ‘unsatisfying taste or smell since it is no longer fresh’, they collocate with different nouns. That is, ‘rancid’ is frequently associated with ‘butter,’ whereas ‘sour’ commonly appears with ‘milk’. It is obvious that substituting synonyms for one another can lead to a loss of naturalness in the language (Crawford & Csomay, 2016, p.40)

In terms of semantic preference, Cheng (2012) pointed out that semantic preference is defined as ‘the possibility for words to be restricted to particular semantic categories. According to Partington (1998), noun collocates of the word ‘sheer’ are classified into four semantic categories depending on their semantic features. The first set includes terms referring to magnitude, weight, or volume, as in ‘the sheer volume of reliable information’ and ‘the sheer weight of noise,’ for example. The second category includes words expressing power, strength, or energy, such as ‘the sheer physical energy’ and ‘the sheer force of an earthquake.’ The third group deals with persistence, as seen by phrases like ‘sheer persistence’ and ‘his sheer indomitability.’ The final one has to do with terms that portray powerful emotions, such as ‘sheer inspiration’ and ‘sheer panic’. However, despite ‘pure’, ‘complete’, and ‘absolute’ having the same meaning, they cannot replace ‘sheer’ in all contexts as they do

not appear to have the same semantic fields as 'sheer'.

### 2.3.5 Grammatical pattern

Investigating a grammatical pattern is another effective criterion for differentiating synonyms. Following Phoocharoensil's (2010, p.5) research, it was discovered that synonyms can have various grammatical patterns. The synonyms 'able' and 'capable', for example, have the same meaning of 'having the qualities or abilities to do something'. They do, however, have different grammatical structures, which means that substituting synonyms results in ungrammaticality. The example, 'able' is followed by an infinitival phrase, as in 'Jonathan is able to fly Concorde,' whereas 'capable,' as in 'Jonathan is capable of flying Concorde,' requires a prepositional phrase.

### Previous studies

Numerous earlier researchers have compared the data derived from corpus software, such as BNC and COCA, to investigate the differences between synonyms.

Similar to this present study, Podzimek (2008) conducted the study with the adjective synonyms 'far', 'distant', and 'remote'. The results of using BNC revealed that, while three words share the meaning of something being far away, they can differ in some semantics and degree of formality. Only 'distant' and 'remote' appear with a noun regarding a relationship and kinship. Solely 'far' can be used to modify a noun concerning a political term, while 'remote' is the only synonym that can collocate with a noun referring to space with a sense of solitary. In addition, in terms of the degree of formality, the results also pointed out that 'remote' is more formal than 'distant' and 'far' since it has the highest frequency in the academic register as compared to 'distant' and 'far', respectively.

Chung (2011) examined the similarities and differences between the synonyms 'create' and 'produce', using the Brown Corpus and the Frown Corpus in combination with the British National Corpus (BNC). The results disclosed that two words appear as bare infinitives and in -ed forms most commonly. The objects that follow the two words, on the other hand, are not the same. 'Produce' is typically followed by

objects which involve mass-produced goods with low creativity, whereas 'create' is typically followed by abstract objects created in smaller quantities with a higher level of creativity. More interestingly, they also have two overlapping meanings. 'Bring into existence or make it happen, occur, or exist' is the first, 'create or manufacture a man-made product' is the second.

The study of Song (2021) confirmed the success of the use of a corpus-based method to help distinguish the two near-synonyms 'damage' and 'destroy', concerning frequencies, the distribution across genres, colligation and collocation, and the differences in meanings and uses. BNC was used to collect data for this study. The findings indicated that 'destroy' and 'damage' are most commonly used in written books and periodicals. However, 'destroy' appears more frequently in spoken and written communication than 'damage.' With regard to colligations and collocations, 'damage' is commonly used in the social science context, whereas 'destroy' is more common in texts related to abstract nouns, namely, belief, thought, arts, and even military affairs. Furthermore, concerning the subtle meanings of the two synonyms, 'destroy' refers to something that no longer exists, in contrast to 'damage', which is used to refer to something that can be recovered but no longer functions as well as earlier.

Hu (2015) conducted the study by employing COCA to investigate the semantic preference and semantic prosody across academic texts, with the three synonymous adjective pairs. The results disclosed that synonyms are not interchangeable as a word seems to be associated with a wide range of semantic properties, identifying various prosodies. Furthermore, the findings show that the three synonymous tend to express neutral meanings in academic texts, supporting the fact that the academic registers normally require neutral meanings. However, the researcher suggested that other researchers may examine data from other academic disciplines and even other registers to obtain more specific details.

Selmistraitis (2020) examined the distribution across different academic texts as well as the

semantic preference and semantic prosody, with the three pair synonyms. The data was drawn from COCA. The findings revealed that near synonyms are varied due to the different semantic preferences as and distribution across registers. For instance, while 'precise' is associated with the words that refer to temporality and locality, 'accurate' collocates with the words in relation to proving, identification, and discovering. Furthermore, in terms of distribution in academic texts, 'accurate' exists more frequently than 'precise.' In addition, regarding semantic prosody, although the synonyms identically express neutral senses, they differ in regard to semantic preferences and the distribution across academic texts, meaning that they cannot be substituted for each other in all contexts.

Petcharat and Phoocharoensil (2017) investigated the meanings, degrees of formality, collocations, and grammatical patterns, based on three adjective synonyms: 'appropriate', 'proper', and 'suited.' The data was gathered from three dictionaries, including COCA. The findings disclosed that the three synonyms have the same basic meaning but vary in the details of their meanings. For example, 'suitable' can be used to describe a person but 'appropriate' cannot. The data also revealed that 'appropriate' appears to have a higher level of formality than 'suitable' and 'proper'. Moreover, in terms of collocations and grammatical patterns, only a few collocates and grammatical structures are shared by the three synonyms. As a result, the three synonyms cannot be used interchangeably in all contexts.

Phoocharoensil (2020) used COCA to investigate the degree of formality and collocations associated with the semantic preference for the three synonyms 'consequence', 'result', and 'outcome'. According to the research, 'result' appears the most frequently in academic texts, followed by 'outcome' and 'consequence', in that order. In addition, in terms of collocation and semantic preference, 'consequence' often co-occurs with words indicating negative meaning, while 'result' expresses the meaning in a positive way rather than 'consequence' and is also dominantly used in research contexts. In

comparison to their synonyms, 'outcome' appears to have the most varied themes.

In another study by Phoocharoensil (2021), the researcher looked at two near-synonyms, 'persist' and 'persevere', using COCA to obtain data. MI scores of at least 3 were used to determine the top 20 noun collocates. The results revealed that 'persist' is more formal than 'persevere' in terms of formal degree. Furthermore, regarding their collocations, the two synonyms have many different noun collocates, resulting in different semantic prosody. 'Persist' is frequently associated with negative words, such as 'stagnant economy', whereas 'persevere' occurs with the context related to Christian, such as 'God' and 'saints', the context of great effort in overcoming a challenge, such as 'marathon' and 'obstacles', and also the context of strong determination, such as 'effort' and 'determined'.

### 3. Methodology

#### 3.1 Data collection and research instruments

The data for this study was derived from the process of COCA for a variety of reasons. First, COCA is one of the most well-known English corpora and is widely used in the field of ELT for research and practice. Second, it is a mega corpus containing over one billion words of text, with over twenty-five million words each year from 1990 to 2019. The texts are also classified into eight genres, namely spoken, fiction, popular magazines, newspapers, academic texts, TV and movie subtitles, blogs, and other web pages, which cover both spoken and written texts. Furthermore, COCA is an annually enlarged corpus resource with new texts (Schmitt, 2010).

#### 3.2 Data analysis

The data was analyzed using two different methods in order to answer the two research questions. First, to investigate the degree of formality, the researcher collected the data of the frequency and also the frequency per million words of use of each word in each genre from COCA, and they were then arranged from the highest to the lowest frequency. It is important to consider frequency compared to a million words since there are many corpus genres and they often have different sizes. Second, the data was drawn from searching collocates through COCA to extract

noun and adverb collocates of the three synonyms in order to see how they differ in collocations and semantic preference. The collocates were determined by MI scores of at least three due to its significance value for the association of collocation (Cheng, 2012). However, Schmitt (2010) recommended that the MI score should be considered with frequency to overcome the limitation of MI value that it may bring rare occurrences to be a collocate. As a result, the collocates were selected from the top 30 frequencies in this study. However, collocates with MI scores of at least three were excluded from this study, supposing that they appear without any relation to that synonym. The target collocates were finally classified following their semantic preference into different semantic categories. Furthermore, the sentences or phrases from the concordance lines were extracted to exemplify how the target synonyms are used in various contexts.

#### 4.Results and Discussion

This chapter presents and discusses the findings in response to research questions pertaining to the degree of formality as determined by the distribution of frequency across genres and also the difference of nouns and adverbs collocating with three target synonyms.

##### 4.1 Degree of formality

As can be seen in Table 2 below, 'far' is most commonly used among their synonyms. since it has the highest frequency of use (333.98 per

million words), which is 14 times higher than 'distant' and 'remote'. The frequency of use of 'distant' (23.42 per million words) is almost identical to that of 'remote' (23.73 per million words). Moreover, 'distant' and 'remote' are also similar in order of frequency of use in the last five genres, namely webpages, newspapers, blogs, TV and movie subtitles, and spoken, respectively.

As for the degree of formality of the three target synonyms, the use of 'far' is most frequent in the genre of blogs, while it is the least frequent in academic texts. This should demonstrate that 'far' is used in informal contexts rather than formal ones. In contrast to 'far', 'distant' appears in the academic text genre as the third order with a frequency of 27.95 per million words, whereas it is least frequent in the genre of spoken language with a frequency of only 4.81 per million words, which should indicate that 'distant' tends to occur more in formal than in informal contexts. In addition to this, 'remote' is similar to 'distant' in that it has the lowest frequency in the same genre with a frequency of 14.32 per million. However, 'remote' appears second in academic texts with a frequency of use of 29.18 per million, which is slightly more than 'distant'. This should point out that 'remote' is used in formal rather than informal contexts and is even more formal than 'distant'. Another indication of a more formal degree of 'remote' is that 'distant' also seems to be used in informal contexts since the most frequent use is found in the genre of fiction (61.01 per million words).

**Table 2: The distribution of the synonyms 'far', 'distant', and 'remote' across eight genres with their frequency of occurrences, derived from COCA**

Rank	'Far'			'Distant'			'Remote'		
	Genre	Frequency	Per million	Genre	Frequency	Per million	Genre	Frequency	Per million
1	Blogs	53,613	416.85	Fiction	7,219	61.01	Magazines	5,049	40.04
2	Fiction	46,490	392.91	Magazines	4,590	36.40	Academic texts	3,496	29.18

Rank	‘Far’			‘Distant’			‘Remote’		
	Genre	Frequency	Per million	Genre	Frequency	Per million	Genre	Frequency	Per million
3	Webpages	47,248	380.25	Academic texts	3,348	27.95	Fiction	3,086	26.08
4	Magazines	45,653	362.06	Webpages	2,438	19.62	Webpages	2,900	23.34
5	Spoken	42,855	339.75	Newspapers	1,843	15.14	Newspapers	2,812	23.10
6	Newspapers	40,029	328.80	Blogs	1,668	12.97	Blogs	2,563	19.93
7	Academic texts	27,462	229.25	TV and movie subtitles	1,289	10.06	TV and movie subtitles	1,853	14.47
8	TV and movie subtitles	28,302	220.98	Spoken	859	6.81	Spoken	1,806	14.32
	Total	331,652	333.98	Total	23,254	23.42	Total	23,565	23.73

The next section presents the results and discussion pertaining to the top 30 noun and adverb collocates of the three target synonyms.

#### 4.2 Collocations and semantic preference

Table 3 illustrates that ‘distant’ and ‘remote’ have the 30 most frequent noun collocates, while ‘far’ has only 13 noun collocates, suggesting that ‘distant’ and ‘remote’ are likely to be used with a broader set of nouns than ‘far’.

In addition, all three synonyms do not have the same noun collocates. Only two of them have the same noun collocates. That is, only ‘distant’ and ‘remote’ co-occur with the nouns, namely ‘galaxy’, ‘location’ and ‘mountain’. As for the synonyms ‘far’ and ‘distant’, they have only two noun collocates in common, i.e. ‘horizon’ and ‘shore’. If compared to their synonyms, ‘corner’ is the only one noun that collocates with ‘far’ and ‘remote’.

**Table 3: Noun collocates of ‘far’, ‘distant’ and ‘remote’ from COCA**

Rank	‘Far’			‘Distant’			‘Remote’		
	Noun collocate	Frequency	MI-Score	Noun collocate	Frequency	MI-Score	Noun collocate	Frequency	MI-Score
1	cry	1792	7.02	future	800	4.85	control	2308	5.90
2	east	1589	4.32	galaxy	761	7.30	area	1485	4.54
3	corner	1137	4.42	past	750	5.46	location	725	5.75

Rank	'Far'			'Distant'			'Remote'		
	Noun collocat e	Frequen cy	MI-Score s	Noun collocat e	Frequen cy	MI-Score s	Noun collocat e	Frequen cy	MI-Score s
4	wall	854	3.05	relative	520	6.73	village	561	5.72
5	distanc e	443	3.26	cousin	512	6.52	sensing	512	10.93
6	left	310	4.95	memory	479	4.67	island	459	4.59
7	shore	254	4.30	star	460	3.88	region	387	4.25
8	horizon	144	3.32	planet	380	4.86	possibilit y	374	4.81
9	fringe	54	3.29	land	365	3.80	site	331	3.15
10	recess	32	3.05	object	363	4.72	corner	328	4.52
11	outback	8	3.20	mountai n	312	4.23	access	290	3.98
12	bulkhea d	7	3.01	second	253	3.20	mountain	256	3.90
13	dais	5	3.05	mile	225	3.29	server	189	5.59
14				horizon	170	5.49	device	184	3.57
15				hill	168	3.57	camera	165	3.10
16				thunder	165	6.68	seller	150	6.06
17				universe	157	4.10	desert	143	4.66
18				shore	144	5.42	outpost	135	7.54
19				metastas is	142	9.29	viewing	131	6.58
20				ancestor	130	5.58	desktop	126	5.94
21				location	121	3.21	valley	119	3.74
22				sky	117	3.17	forest	110	3.20
23				quasar	94	8.66	monitorin g	108	5.68
24				cloud	93	3.40	wildernes s	86	5.12

Rank	'Far'			'Distant'			'Remote'		
	Noun collocates	Frequency	MI-Scores	Noun collocates	Frequency	MI-Scores	Noun collocates	Frequency	MI-Scores
25				echo	87	5.72	province	85	4.66
26				moon	82	3.02	cabin	79	4.43
27				light-year	78	7.08	jungle	77	5.03
28				Siren	74	5.95	sensor	73	4.65
29				supernova	74	6.98	galaxy	72	3.86
30				rumble	72	7.44	observation	70	3.20

The following tables show noun collocates of target words that were placed in their semantic categories based on their semantic preference.

It can be observed from Table 4 that the noun collocates of 'far' were divided into four different categories. That is, the first category refers to far away in regard to elements in space, as shown in the context in (1). The second one denotes distant locations, as exemplified in (2). The next involves a far direction, such as in (3). The last relates to the distance of other issues, namely 'cry', 'fringe' and 'recess', as in the example in (4). Another interesting aspect is that 'far', when collocated with the noun 'cry', expresses not only the direct sense of a far

distance, but also the indirect meaning 'something very different', as seen in (5).

(1) Heat lightning would flare on the **far horizon**.

(2) They pulled themselves across the river, then heaved themselves onto the **far shore**.

(3) In the **far corner** the man was still asleep.

(4) A sudden infusion of light in the **far recess** of the cave.

(5) Our current deficit this year alone is \$1.6 trillion, a **far cry** from the 400 billion we would save annually.

**Table 4: Semantic preference of noun collocates of 'far'**

Semantic Categories	Noun Collocates of 'Far'
1. Spatial elements	horizon
2. Places	bulkhead, dais, outback, shore, wall
3. Direction	corner, distance, east, left
4. Miscellaneous	cry, fringe, recess

As shown in Table 5, the noun collocates of 'distant' fall into seven different categories.

First, it is tied to describe the distance of time, as seen in the example in (6). Second, it is about

distant spatial elements, as observed in (7). Third, it refers to a distant geographical place as well as other locations, as seen in (8). Fourth, it describes a kinship without a close relationship, as exemplified in (9). Next, it is relevant to sound distance, as illustrated in (10). It also relates to the quantity of distance associated with a unit of distance, as in (11). Finally, it regards the distance of other issues, namely ‘memory’, ‘metastasis’, ‘object’ and ‘second’, as demonstrated in (12). Apart from these findings, a more interesting finding is the position of ‘distant’. ‘Distant’ normally precedes the noun that it modifies, as in (13), in contrast, when collocating with a noun that represents a unit of a distance, i.e., ‘light-year’ and ‘mile’, it is inserted after that noun, as seen in (14).

(6) In the **distant future** it is a certainty that the Earth will warm beyond what we have today.

(7) The Hubble Space Telescope continued to find surprising objects in the **distant universe**.

(8) The sun's rim had just appeared between two **distant hills**.

(9) There may be many **distant relatives** naturally dependent on our aid.

(10) Her voice sounded far away, like a **distant echo** in the woods.

(11) The former locality is 50 **miles distant** in a straight line from the outcrop of tar-sand at Boiler Rapid.

(12) Those things bothered me at the time, but now they're **distant memories**.

(13) We are **distant cousins**.

(14) The latter is 106 **miles distant** from the same point.

**Table 5: Semantic preference of noun collocates of ‘distant’**

Semantic Categories	Noun Collocates of ‘Distant’
1. Time	future, past
2. Spatial elements	cloud, galaxy, horizon, moon, planet, quasar, sky, star, supernova, universe
3. Places	hill, land, location, mountain, shore
4. Kinship	ancestor, cousin, relative
5. Sound	echo, rumble, siren, thunder
6. Measurement of distance	light-year, mile
7. Miscellaneous	memory, metastasis, object, second

The noun collocates of ‘remote’ were separated into six different groups, as shown in Table 6. The first group refers to a distant spatial element, as seen in (15). The second one denotes distant places, as exemplified in (16). The third one identifies the distance of direction, as observed in (17). The fourth one refers to electronic devices that allow a user to connect to from a distant location, as demonstrated in (18). The next group, as shown

in (19), pertains to computer systems and servers available for users to access a computer or device from another device from anywhere. The last one refers to the distance of other concerns, i.e. ‘observation’, ‘possibility’, ‘seller’, ‘sensing’, and ‘viewing’, as seen in (20). Another intriguing finding is that when ‘remote’ collocates with a noun representing electronic devices or even computer systems and servers, it always precedes that noun and

appear together as a compound noun, suggesting that ‘remote’ followed by those nouns is extensively used as a compound noun, as illustrated in (21).

(15) JWST also will be able to view even more **remote galaxies**.

(16) John buries the corpses in a **remote forest**.

(17) There is a small village graveyard in one of the **remote corners** of Russia.

(18) It was a **remote device** that could transmit the images.

(19) Applications written in HTML often require mechanisms to communicate with **remote servers**.

(20) Notice requirements are typically aimed at the **remote seller**.

(21) The **remote desktop** is displayed in a window on the desktop.

**Table 6: Semantic preference of noun collocates of ‘remote’**

Semantic Categories	Noun Collocates of ‘Remote’
1. Spatial element	galaxy
2. Places	area, cabin, desert, forest, island, jungle, location, mountain, outpost, province, region, site, valley, village, wilderness
3. Direction	corner
4. Electronic equipment	camera, control, device, sensor
5. Computer systems / servers	access, desktop, monitoring, server
6. Miscellaneous	observation, possibility, seller, sensing, viewing

Regarding the semantic categories to which most of the noun collocates of each target word belong, the findings indicate that ‘far’ and ‘remote’ are most commonly used to describe the distance between places, in contrast, ‘distant’ is most frequently used to express the distance of spatial elements. What's more, the data shows that all synonyms have only two semantic categories in common: spatial elements and places, implying that synonyms can have different deeper meanings. Moreover, this finding also backs up Podzimek’s (2008) conclusion that three words would have some semantic differences. According to the findings, some synonyms can be used in particular contexts. To demonstrate this point, only ‘distant’ is used to denote the meaning of a noun regarding time, kinship, sound, as well as the measurement of distance, while only ‘remote’ is used to describe far away in terms of electronic equipment and computer systems. Furthermore, only ‘far’ and ‘remote’ are

encountered when referring to distant directions.

In the next section, the researcher proceeds to provide the findings and discussion regarding the adverb collocates of the target synonyms.

As can be seen in Table 7, there are no adverb collocations of ‘far’ needed to be examined in this study. More specifically, although an adverb, like ‘thus’, possesses a MI-score of more than 3, it is not considered an adverb collocate, since MI-score often draws the rare occurrence of a word as a collocate. As shown in (22), the word ‘thus’ does not appear with the word ‘far’, but rather with the word ‘further,’ which just has a similar meaning to ‘far.’ In addition, Table 6 also demonstrates that ‘distant’ and ‘remote’ share only four adverb collocates, i.e., ‘geographically’, ‘impossibly’, ‘unimaginably’ and ‘temporally’. Moreover, the number of adverb collocates of ‘distant’ (29

collocates) is about two times higher than that of 'remote' (8 collocates), indicating that 'distant' is used with adverbs more commonly than 'remote'.

(22) **Thus**, further study is needed to determine the potential retinal effects of domestic lighting with high blue.

**Table 7: Adverb collocates of 'far', 'distant' and 'remote' from COCA**

Rank	'Far'			'Distant'			'Remote'		
	Adverb collocates	Frequency	MI-Scores	Adverb collocates	Frequency	MI-Scores	Adverb collocates	Frequency	MI-Scores
1				emotionally	83	5.12	geographically	26	5.80
2				increasingly	70	3.05	exceedingly	9	3.94
3				somewhat	63	3.06	impossibly	9	4.47
4				geographically	39	6.43	unimaginably	6	6.50
5				seemingly	29	3.03	temporally	5	5.44
6				impossibly	23	5.87	horizontally	3	3.06
7				strangely	19	4.02	hauntingly	2	4.87
8				oddly	18	3.80	unfathomably	2	6.80
9				culturally	16	3.43			
10				infinitely	9	3.87			
11				ideologically	7	4.34			
12				temporally	7	5.97			
13				unimaginably	7	6.76			
14				phylogenetically	6	8.19			
15				comparatively	5	3.08			
16				weirdly	3	3.82			

Rank	‘Far’			‘Distant’			‘Remote’		
	Adverb collocates	Frequency	MI-Scores	Adverb collocates	Frequency	MI-Scores	Adverb collocates	Frequency	MI-Scores
17				spatially	3	4.08			
18				immeasurably	3	4.90			
19				evolutionarily	3	5.46			
20				maximally	3	5.52			
21				fantastically	2	3.84			
22				chronologically	2	4.00			
23				frighteningly	2	4.12			
24				reverently	2	4.50			
25				cordially	2	4.93			
26				irretrievably	2	5.16			
27				experientially	2	6.09			
28				inconceivably	2	6.68			
29				epistemically	2	6.97			

In the following tables, only semantic preferences of adverb collocates of ‘distant’ and ‘remote’ are demonstrated.

As can be seen in Table 8, the adverb collocates of ‘distant’ were divided into five semantic groups. The first group describes how far a place is, as in (23). The following groups refer to the distance in terms of emotion, geography, and time, as exemplified in (24), (25), and (26), respectively. The last group pertains to the distance in relation to other issues, such as ‘culturally’, ‘experientially’ and ‘ideologically’, as demonstrated in (27). Another interesting result is that, with the

exception of the word ‘somewhat’, almost all the adverbs collocates of ‘distant’ are the adverbs formed from adjectives integrated with the suffix ‘-ly’, such as ‘infinitely’ and ‘fantastically’. Furthermore, ‘distant’ occurs most often with adverbs that are considered a type of adverb of degree and it is placed before the modified adjective ‘distant’ to convey how far away something is, as in (28).

(23) It can only be found in a place **infinitely distant** from this world.

(24) These things are in **fantastically distant** galaxies.

(25) Emotional sites may be in **geographically distant** places.

(26) ...As transposing the spatially distant and timeless generic village to a **temporally distant** collective past.

(27) Even younger children seem somehow more **culturally distant** from their parents.

(28) It can only be found in a place **infinitely distant** from this world.

**Table 8: Semantic preference of adverb collocates of ‘distant’**

Semantic Categories	Adverb Collocates of ‘Distant’
1. Amount of distance	comparatively, immeasurably, impossibly, inconceivably, increasingly, infinitely, maximally, somewhat, unimaginably,
2. Emotion	cordially, emotionally, fantastically, frighteningly, irretrievably, oddly, strangely, weirdly
3. Geography	geographically, spatially
4. Time	chronologically, evolutionarily, temporally
5. Miscellaneous	culturally, epistemically, experientially, ideologically, phylogenetically, reverently, seemingly

Table 9 shows how the adverb collocates of ‘remote’ were classified into five semantic categories. As observed in the first group, it describes the amount of distance of a place, as seen in (29). The following categories, as shown in (30), (31), (32) and (33), refer to the distance in relation to emotion, geography, time, and direction, respectively. Furthermore, similar to ‘distant’, the adverbs collocated with ‘remote’ mostly appear with the adverb of degree, as exemplified in (34). Those adverbs are formed in the same way as the adverb collocates of ‘distant’.

(29) OCS waters are **exceedingly remote**.

(30) An army explorer stumbled onto a **hauntingly remote** and desolate chunk of Texas prairie.

(31) Big oil companies have turned to places that are **geographically remote** and politically volatile.

(32) That challenges us from a **temporally remote** and culturally estranged past.

(33) In many myths the fire is stolen from a **horizontally remote** location.

(34) They happen every second or so, usually in some **unimaginably remote** galaxy.

**Table 9: Semantic preference of adverb collocates of ‘remote’**

Semantic Categories	Adverb Collocates of ‘Remote’
1. Amount of distance	exceedingly, impossibly, unfathomably, unimaginably
2. Emotion	hauntingly
3. Geography	geographically

<b>4. Time</b>	temporally
<b>5. Direction</b>	horizontally

It is, therefore, in terms of adverb collocates of ‘distant’ and ‘remote’, they mostly occur with the adverbs under the category pertaining to the amount of distance. They also share four semantic preference, i.e. the amount of a distance and the distance in relation to emotion, geography, and time. However, they do not share all contexts. That is, only ‘remote’ appears with the adverbs referring to direction, whereas ‘distant’ appears to be used instead of ‘remote’ when associating with the adverbs relating to other issues, such as ‘culturally’, ‘experientially’, so on.

To conclude, the findings of this study support Podzimek (2008) whose study identified that ‘remote’ has the highest formal degree, followed by ‘distant’, and ‘far’ respectively. As a result, some synonyms can differ in terms of formal degree (Nisani, 2016; Petcharat & Phoocharoensil, 2017; Phoocharoensil, 2020; Ruenroeng, 2014.). That is, one synonym is used in a more formal context than another (Jackson & Amvela, 2000). Evident from the findings concerning the collocations and semantic preference, it appears that the target synonyms do not share all the same noun or adverb collocates, supporting the fact that a synonym is more likely to appear alongside a specific word than with others (Lehecka, 2015). They also prefer some different semantic categories and are not evenly distributed between categories, which are in line with the conclusion of the study by Selmistraitis (2020). Another interesting finding is that, contrast to their synonyms, ‘far’ rarely collocates with an adverb. Furthermore, the current findings of this study are consistent with numerous previous studies stating that synonymous words show some distinctions in terms of collocation and semantic preference (Phoocharoensil, 2021; Selmistraitis, 2020; Song, 2021). Thus, although the three adjective words are synonyms, they differ in terms of formality, collocations and semantic preference, and hence cannot be used interchangeably in all contexts (Cai, 2012; Hu, 2015; Phitayakorn, 2016; Phoocharoensil, 2010).

## 5. Conclusion

As the findings point out, although the adjective synonyms ‘far’, ‘distant’ and ‘remote’ are similar in some contexts, they cannot be absolutely substituted in every context due to the difference in formality and collocations. Despite the fact that they seem to occur in both formal and informal contexts, they are more likely to be used in either context. Moreover, as for their collocations, there are only few noun collocates shared by the three synonyms. Similarly, only a few adverb collocates are shared by ‘distant’ and ‘remote’. In addition, in terms of semantic preference, the differences between the other semantic preference that are not shared seem to indicate the subtle variation in meaning between the three synonyms.

Consequently, because of the findings of this study, it could be beneficial for English teachers. Clearly, this research provides teachers with an authentic resource for finding examples or creating materials to help students comprehend the differences between these target synonyms and then use them correctly and naturally in different contexts. Furthermore, because of the success of COCA in recognizing near-synonyms, this study also provides teachers with the opportunity to create authentic content using corpus data. According to Johns (1991), this type of content encourages students to engage in Data-Driven Learning (DDL), which supports students to learn language through autonomous learning, inductive learning, and even cognitive skills, all of which are important for language learners.

The current study, however, has some limitations. First, this study primarily relied on the use of COCA to collect data. Because of that, future studies could employ other well-known native-speaker corpora, such as the British National Corpus (BNC), to see whether there are more interesting results to be found. Second, the researcher of this study conducted the study with emphasis on two criteria. It may be more interesting to analyze synonyms by using other criteria, such as dialects and grammatical patterns.

## References

1. Barnbrook, G, Mason, O. & Krishnamurthy, R. (2013). In *Collocation: Applications and Implications*. Palgrave Macmillan.
2. Cai, J. (2012). Is it 'great' enough? A corpus-based study of 'great' and its near synonyms. Indiana: Ball state University.
3. Cheng, W. (2012). Exploring corpus linguistics. *Language in action*. Routledge.
4. Chung, S. F. (2011). A corpus-based analysis of 'create' and 'produce'. *Chang Gung Journal of Humanities and Social Sciences*, 4(2), 399-425.
5. Crawford, W. J. & Csomay, E. (2016). *Doing corpus linguistics*. Routledge.
6. Flowerdew, L. (2012). *Corpora and language education*. Palgrave Macmillan.
7. 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.
8. Jackson, H., & Amvela, E. (2000). *Words, meaning and vocabulary: Meaning relations*. Athenaum Press.
9. Lehecka, T. (2015). *Collocation and colligation*. In *Handbook of pragmatics online*. Benjamins.
10. Lewis, M. (1997). *Implementing the lexical approach*. Language Teaching Publications.
11. Longman Dictionary of Contemporary English. (2014). Essex: Pearson Education.
12. McCarthy, M. and O'Dell, F. (2005). *English collocations in use*. Cambridge University Press.
13. Nelson, M. (2006). Semantic association in business English: A corpus-based analysis *English for Specific Purposes*, 25, 217–34. <https://doi.org/10.1016/j.esp.2005.02.008>.
14. Nisani, F. (2016). A Corpus-based Study of English Synonyms: Possible, Probable, and Likely. Language Institute, Thammasat University.
15. Palmer, F. R. (1997). *Semantics*. Great Britain: Cambridge University Press.
16. Partington, A. (1998). *Patterns and meanings. Using corpora of English language research and teaching*. John Benjamins.
17. Partington, A. (2004). 'Utterly content in each other's company': Semantic prosody and semantic preference. *International Journal of Corpus Linguistics*, 9(1), 131-156.
18. 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.
19. Phitiyakorn, M. W. (2016). A Corpus-based Study of English Synonyms: Advise, Recommend, and Suggest. (Independent Study, Thammasat University).
20. Phoocharoensil, S. (2010). A corpus-based study of English synonyms. *International Journal of Arts and Sciences*, 3(10), 227-245.
21. Phoocharoensil, S. (2020). A Genre and Collocational Analysis of Consequence, Result, and Outcome. 3L: Southeast Asian Journal of English Language Studies, 26(3).
22. 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.
23. Podzimek, J. (2008). *On Conceptual Synonymy (A Case Study)* (Doctoral dissertation, Masarykova univerzita, Filozofická fakulta).
24. Ruenroeng, C. (2014). A corpus-based analysis of English synonyms: ruin, demolish, destroy. Bangkok, Thammasat university.
25. Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Springer.
26. Selmistraitis, L. (2020). Semantic Preference, Prosody and Distribution of Synonymous Adjectives in COCA. *GEMA Online Journal of Language Studies*, 20(3), 1-18.
27. Song, Q. (2021). Effectiveness of Corpus in Distinguishing Two Near-Synonymous Verbs: 'Damage' and 'Destroy'. *English Language Teaching*, 14(7), 8-20.
28. Webb, S. & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.