

Are You Being *Naughty*, *Disobedient*, or *Rebellious*? A Corpus-Based Study of English Synonyms

Received : Aug 31, 2021

Revised : Oct 18, 2021

Accepted : Nov 30, 2021

Punyapa Boontam

Division of Foreign Languages,
Nakhon Sawan Rajabhat University

Abstract

English synonyms can be used to expand the vocabulary skills of students; however, students tend to encounter thesaurus errors when they apply synonyms in the wrong context. Therefore, in order to differentiate synonyms effectively, corpus-based analysis should be applied. This research study examined the similarities and differences between *naughty*, *rebellious*, and *disobedient* in terms of meanings, collocations, degrees of formality, and grammatical patterns across the different standards of English, namely British and American English. The corpus data was drawn from the British National Corpus (BNC) with the assistance of the online tool Sketch Engine (SkE) and the Corpus of Contemporary American English (COCA) to represent different varieties of English which can broaden the knowledge of the students across cultures. The results show that these three synonyms are taken as loose synonyms because they cannot be used interchangeably in all contexts even though they could share the same core meaning as an act of not being willing to follow or obey someone or something. The corpus data from BNC and COCA reveal some shared characteristics and contrastive features. Therefore, the findings of the study may have a profound impact in the fields of English language teaching, especially in a digital technology-driven age, as corpus data can serve as a useful tool to explore the authentic language patterns and help enhance learners' metalinguistic awareness.

Keywords: Corpus-based Analysis, Corpus Linguistics, English Synonyms, Vocabulary Teaching

1. Introduction

In a digital technology-driven age, it is undeniable to say that learning English is far more convenient and accessible for ESL/EFL learners around the world. Digital technology, which is both available online and offline, could enable students to have more exposure to a number of English learning resources. Precisely, English language skills play a vital role in the 21st century and can help students achieve their goals and lead to career opportunities as English serves as a tool that enables them to communicate effectively with people from other countries (Kirkpatrick, 2007). Nevertheless, among the four main skills of English, vocabulary is considered as the most essential skill because lexical competence and lexical knowledge are regarded as the foundation for learners to comprehend both written and spoken languages and

lead to successful communication (Alquahtani, 2015; Ling, 2005; Nation, 1994). In the same vein, the lack of vocabulary knowledge can limit the number of sentences that they would like to produce and consequently lead to learners' frustration (Caro & Mendinueta, 2017). Considering this problem, ESL/EFL learners tend to experience the problem of lexical variety when they are assigned to do the writing task in English (Ferris, 2009). Despite its importance, English is implicitly considered as one of the most difficult subjects in Thailand as students tend to encounter difficulties when learning English, especially when it comes to real-life communication due to their limited vocabulary knowledge (Sasum & Weeks, 2018).

Nonetheless, one of the effective ways to expand the students' vocabulary skills is using synonyms as investigating synonyms can help enhance students' metalinguistic knowledge, for example, part of speech, register, and collocations (Shaw, 2011). Sotoudehnama and Soleimanifard (2013) found that teaching new words through synonyms could enhance the students' short-term vocabulary learning and increase the excitement in the classroom. Moreover, the study of synonyms plays an important role in terms of improving learners' accuracy and fluency in English (Shahzadi et al., 2019). Learning synonyms is also crucial for learners who seek advancement in writing as they would be able to use a variety of words in their work (Phoocharoensil, 2010). Liu (2010) also explained that learners with the ability to use various synonyms would be able to convey their message effectively and rarely fail to interpret the meaning. However, there is a concern when teaching synonyms to students known as *thesaurus errors* which tend to occur when learners apply synonyms in the wrong context (Ferris, 2009). Due to L1 interference, differentiating meanings and usages of synonyms poses a difficulty for most EFL/ESL learners as translating synonyms directly into their L1 might lead to confusion and thesaurus errors (Aroonmanakun, 2015). Significantly, one factor that ESL/EFL teachers and learners should be aware of when using English synonyms is that some synonyms cannot be used interchangeably in some contexts. It should be noted that relying solely on intuition and definitions provided by dictionaries is not sufficient to distinguish synonym pairs (Francis et al., 1996; Phoocharoensil 2020). Thus, corpus data can provide valuable insights into naturally occurring language that is largely unavailable to linguistic intuition (Sinclair, 1991; Szudarski, 2018) and serve as a reliable tool to distinguish synonyms effectively (Biber et al., 1998; Hunston, 2002; Moon, 2010).

This research study aims to examine the similarities and differences between *naughty*, *rebellious*, and *disobedient* in terms of collocational patterns, degrees of formality, and grammatical patterns across the different standards of English, namely British and American English through a corpus-based investigation. The selection of the synonyms was determined based on their appearance in the British National Corpus (BNC)/Corpus of Contemporary American English (COCA) headword lists (Nation, 2017). The BNC/COCA lists are designed primarily for learners of English as a foreign language which contains the headwords from the 25,000 BNC/COCA word families that come with the Range program. Specifically, *naughty* was listed in the first 1,000 BNC/COCA headwords, and *disobedient* was listed in the eighth 1,000 BNC/COCA headwords, reflecting their high frequency of occurrence in real-life language use. Significantly, *rebellious* was listed in the Academic Vocabulary List (Gardner & Davies, 2014) which contains 20,000 or so most frequent words (lemmas) in the 120 million words of COCA Academic. Consequently, according to the lists, the high frequency of occurrence of each target synonym indicates its usefulness for EFL learners and their teaching worth. Furthermore, these three synonyms can confuse EFL learners because of their overlapping meaning in the L1 language. By way of illustration, there are many different shades of meanings and wordplay in Thai language that could refer to *naughty*, *disobedient*,

and *rebellious*. For example, *deuu*(^{ดื้อ}) in Thai language could be translated into *naughty* and *disobedient*. However, there is a variety of *deuu* in Thai wordplays such as *deuu dann*, *deuu dan*, *deuu deung*, and *deuu phaaeng* that might pose a difficulty for Thai students as it can be challenging for low proficient learners to notice the differences of these synonyms when translating from Thai into English. In this case, *deuu phaaeng* cannot be translated into *naughty* nor *disobedient*, but it should be translated as *rebellious*. Therefore, the results of this study would be beneficial for EFL/ESL teachers as it helps shed light across L2 education and raise awareness among teachers of the implementation of a corpus in teaching synonyms to EFL/ESL students which is suitable for today's data-driven age because it can enhance both academic knowledge and ICT skills at the same time.

2. Literature Review

2.1 Corpus Linguistics

O'Keeffe et al. (2007) have defined a corpus as "a collection of texts, written or spoken, which is stored on a computer" (p. 1). Therefore, corpus linguistics is the study of language through investigating the authentic data in a naturally-occurring context which is beneficial for language learners in terms of authenticity, speed, and reliability (Lindquist, 2009; Sinclair, 1991).

\$1 per share. The company also granted Phemus warrants to	buy	350,000 shares of common stock at \$1 a share . # DELRINA
4 years ago. The announcement that Saudi Arabia is to	buy	48 of the aircraft has come as a relief to workers at
8,928.55 +7,500= 16,428.55 to invest. He or she may now	buy	5 per cent of the equity of the ungeared firm , ABC
, of Britain 's 100 most prescribed drugs , he can	buy	67 abroad . sell them for 20 per cent less than the
, exclusive voucher with you. Hand it over when you	buy	a big value Shows'n'Rides sheet of 12 individual tickets and you
I 'd improved a little bit. And erm managed to	buy	a car and I paid four hundred and ten for it .
the vodka (pause) and say it was n't even enough to	buy	a carry out (pause) but do n't say it unless she says
set their own limits. If you can only afford to	buy	a certain amount of organic produce , potatoes would be a good
you feed your pet on unnecessarily expensive food or could you	buy	a cheaper alternative .? If you grow houseplants try propagating
wife Shirley -- who have a married daughter -- will also	buy	a few things for the house with their summer windfall . #
it work ? Is this safe ? # 3) #	Buy	a fuzz box . I have tried the new Marshall Bluesbreaker FX
to come on Central News. why Americans are flocking to	buy	a great British boot # Welcome back . A down to earth
? ' Denis protested. ' Well , it wo n't	buy	a gun . Perhaps a few bullets ? But give it to
Miranda , not many people of your age can afford to	buy	a house like this . KITS must be doing frightfully well .
, but with a three-year contract , he is planning to	buy	a house . ' Perhaps in Bearsden or Newton Mearns , '

Figure 1. KWIC Concordance lines of *buy* in BNC

In order to explore the corpus data, concordance lines, a list of sentences showing an occurrence of the target word in a particular context (Lindquist, 2009), need to be analyzed systematically. The search word or phrase is called the 'node' which is usually presented in the center of the concordance line known as Key-Word-In-Context displays (KWIC concordances), as shown in Figure 1. KWIC concordance lines will help enhance learners' awareness through investigating the lexico-grammatical patterns by looking at words or phrases which precede or follow the node word. From Figure 1, the right of the node word shows that *buy* is usually followed by an object noun introduced by article *a*, indicating a high instance of noun phrases which is also post-modified by prepositional phrase introduced by *of* and *like*. On the other hand, the left of the node word shows that *buy* is usually preceded by preposition *to*, suggesting prepositional phrase. Likewise, learners can investigate the most frequent collocations of the target words which occur in a specific context by looking at the

frequency figures. It should be noted that the most basic statistical measure in corpus linguistics is a frequency count which is the number of instances or hits of an item that occurs in a corpus.

	CONTEXT	FREQ	ALL	%	MI
1	MONEY	344	36031	0.95	3.25
2	SELL	197	7455	2.64	4.72
3	SHARES	195	8257	2.36	4.56
4	BUY	154	12084	1.27	3.67
5	AFFORD	144	4347	3.31	5.04
6	GOODS	127	9928	1.28	3.67
7	CLOTHES	102	6858	1.49	3.89
8	EQUIPMENT	94	8716	1.08	3.42
9	EXPENSIVE	70	5666	1.24	3.62
10	DRINK	69	7627	0.90	3.17

Figure 2. Frequency figures of *buy* in BNC.

From Figure 2, the most frequent words that co-occur with the node word 'buy' are *money*, *sell*, *shares*, *buy*, *afford*, *goods*, *clothes*, *equipment*, *expensive*, and *drink*, respectively. It can be interpreted that the collocation of *buy* is associated with business terms, such as *money*, *shares*, and *goods*. Also, other possible verbs that could occur with *buy* are *sell* and *afford*, whereas *expensive* is the most frequent adjective that co-occurs with *buy*. These findings could help EFL/ESL learners to understand the contextual usage of the word effectively and lead them to the further investigation of language patterns as this could shed light on other new words or grammar rules that they have never experienced before (Jones & Waller, 2015).

2.2 Corpus-based Study of English Synonyms

Oxford Advanced Learner's Dictionary (2005) defines a *synonym* as “a word or expression that has the same or nearly the same meaning as another in the same language” (p. 1557), such as *big* and *large*. Synonyms can be classified into strict or absolute synonyms and loose or near-synonyms based on the degree of overlapping use of the words in a particular context. Regarding strict or absolute synonyms, two words can be used interchangeably, meaning that one could substitute the other with no effects on meanings and language patterns. In contrast, loose or near-synonyms may share some overlapping meanings and they cannot replace each other in some contexts (Jackson & Amvela, 2000). Their overlapping meanings may differ to a greater or lesser degree in certain areas (Jackson & Amvela, 2007). For example, the synonyms *discover* and *find* can be used interchangeably in the context *She discovered/found that her husband was having an affair*, but not in the context *She doesn't find it easy to talk about her problems*, or in *Who discovered America?*. Hence, loose synonyms are commonly found in English whereas strict synonyms are almost non-existent (Palmer, 1997).

More importantly, Jackson and Amvela (2000) posit some useful criteria that can be used to distinguish near-synonyms, namely dialects and standards of English, formality or style, and connotation. To elaborate, standards of English such as British, American, or Australian English can shed some light on differences between synonyms; for example, *dummy* (BrE and AusE) and *pacifier* (AmE), and *port* (AusE), *luggage* (BrE), and *baggage* (AmE). Moreover, formality or style should be taken into consideration when differentiating synonyms, for example, *book* (informal) and *reserve* (formal), and *ask* (informal) and *enquire* (formal). Lastly, in order to determine any differences between synonym pairs, connotations concerning negative or positive senses should be pointed out; for instance, *famous* (positive)

and *notorious* (negative), and *impulsive* (positive) and *impetuous* (negative). Thus, corpus-based analysis can be considered as an effective tool to analyze the synonyms systematically according to register difference, frequency, and collocations (Biber et al., 1998; Moon, 2010).

Chung (2011) found that synonyms *create* and *produce* shared some senses of meaning, but the corpus data drawn from Brown Corpus and Frown Corpus revealed that *create* seemed to co-occur with objects in a fewer quantity which required a high level of creativity, whereas *produce* tended to be followed by objects in a larger quantity with a low level of creativity. Jarunwaraphan and Mallikamas (2020) also arrived at the same conclusion that near-synonyms can have different connotations despite their overlapping core meanings. With the assistance of COCA, it was discovered that the verb collocates of *chance* tended to convey negativity (e.g., *decrease chances* and *lost all chances*), on the other hand, the collocates of *opportunity* revealed more positive connotations (e.g., *creates enormous opportunities* and *afforded the opportunities*).

Considering usage of synonyms, knowing the register of the words is helpful for students to distinguish the different usages of the synonyms, such as the differences between spoken and written language (Shaw, 2011). For example, *shot* in the academic register could refer to a medical injection; however, when it occurs in fiction or spoken context, it employs the meaning of a gunshot. Implicitly, corpus data could reveal the register difference by examining the frequency of the words that occurred in different text genres, such as spoken, fiction, magazine, newspaper, and academic.

Dealing with genre differences, the degree of formality could be revealed. For example, if a word is frequently used in an academic context, it can be indicated that that word mostly occurs in a formal context. Cai (2012) found that *great* was used more frequently than the other six synonymous adjectives: *awesome*, *excellent*, *fabulous*, *fantastic*, *terrific*, and *wonderful*, obtained from COCA corpus. Interestingly, it is evident that *fabulous*, *fantastic*, *great*, *terrific*, and *wonderful* occurred in spoken genre, whereas *awesome* and *excellent* were commonly used in magazines. The findings suggested that these synonyms were common in an informal context. Using COCA, Phoocharoensil (2020) also found that *error* had the highest tendency to occur in academic texts, reflecting the high degree of formality, whereas *mistake* and *fault* mostly occurred in the informal contexts; such as television and film subtitles.

In terms of collocations, different collocates that co-exist with the target synonyms can implicitly identify the ability of a set of synonyms to substitute each other in some particular contexts. Sumintana and Tawilapakul (2019) investigated the synonyms *cure* and *heal*, using the corpus data obtained from COCA. It was found that *cure* is frequently followed by disease and problems such as *cancer* and *disease*, while *heal* frequently co-occurs with emotional injuries and something abstract, for example, *wounds* and *people*. Hence, they are loose synonyms due to differences in collocational patterns. In a similar vein, according to the corpus data in COCA, Aroonmanakun (2015) found that *fast* was used to emphasise the high speed of the action (e.g., *fast bike*, *fast computers*, *fast internet*), on the other hand, *quick* was often associated with the action being carried out in a short time (e.g., *quick answer*, *quick call*, and *quick drink*). Similarly, corpus data can also increase students' collocational awareness as it reveals the frequency of collocations of the words and provides a mutual information (MI) score which shows how frequently each word is connected (Shaw, 2011). According to Hunston (2002), an MI score of 3 or higher indicates that two items are collocated. Therefore, apart from looking at the frequency, MI scores should also be taken into consideration in order to see the collocational strength of the words.

Lastly, grammatical patterns can unveil the similar or different syntactic patterns of the synonyms which can help to distinguish the strict and loose synonyms. Jirananthiporn (2018) also suggests that the learners should be provided with the lexico-grammatical patterns when synonyms are introduced. Gu (2017) examined the synonyms *obtain* and *gain* through BNC corpus. They found that the syntactic patterns of these two words are similar in the pattern of “verb + object noun” and “verb + adverb”. Nonetheless, the result shows that *obtain* is more frequently used in the passive voice. By the same token, Petcharat and Phoocharoensil (2017) found that *proper*, *appropriate*, and *suitable* cannot be used interchangeably in all grammatical contexts despite their shared common core meaning.

3. Research Methodology

This study is a contrastive corpus-based synonymous analysis. Significantly, a “corpus-based” investigation is often undertaken to check the researcher’s intuition about language use which is distinctive from a “corpus-driven” investigation where the researcher approaches the corpus data with an open mind to explore the emerged patterns (Tognini-Bonelli, 2001). Considering the scope of the study, a corpus-based analysis was employed to investigate the collocational patterns, formality, and grammatical patterns of *naughty*, *rebellious*, and *disobedient* to confirm their status as near-synonyms.

3.1 Sample

Three synonyms *naughty*, *rebellious*, and *disobedient* were selected to be analyzed in this present study, following the criteria of collocations, formality, and grammatical patterns. The selection of the synonyms was determined based on their appearance in the BNC/COCA Level word family lists, namely 1st 1,000 BNC/COCA headwords (Nation, 2017a) and 8th 1,000 BNC/COCA headwords (Nation, 2017b). The BNC/COCA lists are designed primarily for learners of English as a foreign language which contains the headwords from the 25,000 BNC/COCA word families that come with the Range program. Specifically, *naughty* was listed in the first 1,000 BNC/COCA headwords, and *disobedient* was listed in the eighth 1,000 BNC/COCA headwords, reflecting their high frequency of occurrence in real-life language use.

Table 1 shows the definitions of *naughty*, *rebellious*, and *disobedient* from the Cambridge University Press (n.d.).

Table 1. Definitions of *naughty*, *rebellious*, and *disobedient* in Cambridge University Press (n.d.)

<i>Synonyms</i>	<i>Definitions</i>
<i>Naughty</i>	<ol style="list-style-type: none"> 1. When children are naughty, or their behaviour is naughty, they behave badly or do not do what they are told to do 2. Used slightly humorously to describe an adult who has behaved badly or an adult's bad action 3. Involving or suggesting sex

<i>Synonyms</i>	<i>Definitions</i>
<i>Rebellious</i>	<ol style="list-style-type: none"> 1. If a group of people is rebellious, they oppose the ideas of the people in authority and plan to change the system, often using force 2. If someone is rebellious, they are difficult to control and do not behave in the way that is expected
<i>Disobedient</i>	<ol style="list-style-type: none"> 1. Refusing to do what someone in authority tells you to do 2. Not doing what you are told to do

Given the definitions, *naughty*, *rebellious*, and *disobedient* share the core meaning suggesting the action of resistance and the sense of not doing what one is told or expected to do. However, there are some overlapping concepts shared by the three synonyms which are not explicitly explained in the dictionary. Consequently, the choice of the synonyms was predetermined by the fact that *naughty*, *rebellious*, and *disobedient* appear to be problematic and confusing among EFL learners due to their similar L1 meanings. Therefore, a corpus-based analysis should be used as evidence to broaden the understanding of the similarities and differences of the near-synonyms and provide insightful information on unique patterns and usages in naturally-occurring settings.

3.2 Data Collection Procedure

The analysis was conducted in both quantitative and qualitative ways with the assistance of the online tool Sketch Engine (SkE). Sketch Engine is a useful and convenient toolbox to elicit the data from BNC corpus as it provides KWIC concordance lines, frequency/tokens, collocation, word sketch, sketch differences, and syntactic patterns/colligation. In order to ensure the representativeness of the data, a well-balanced corpus containing a variety of subsections representing different genres of language use should be utilized (Hunston, 2002). Therefore, corpus data used in this study were drawn from two significant corpora, namely BNC representing British English, and COCA representing American English. To elaborate, BNC was originally created by Oxford University Press in the 1980s and early 1990s, containing 100 million words of text from a wide range of genres (e.g., spoken, fiction, magazines, newspapers, and academic). Moreover, in order to triangulate the data, the corpus data drawn from COCA, one of the largest American English corpora and the counterpart to BNC, was collected to investigate the dispute of the three synonyms across different varieties of English. Specifically, COCA is composed of more than one billion words in 485,202 texts, including 24-25 million words added each year from 1990-2019 which is evenly divided between the genres of TV and Movies subtitles, spoken, fiction, popular magazines, newspapers, and academic journals. Therefore, BNC and COCA are the only large, well-balanced corpora of English that are freely available online which can complement each other. In terms of genre balance, BNC is 10% spoken and 90% written, while in COCA the corpus is nearly evenly divided (20% in each genre) between spoken, fiction, popular magazines, newspaper, and academic. BNC has a much wider range of spoken sub-genres, while COCA is composed of unscripted conversations on television and radio shows. Both corpora are very well balanced in terms of sub-genres for the written genres (e.g., Newspaper-Sports, or Academic-Medicine). However, due to a discrepancy in genre balance between COCA and BNC, only the corpus data drawn from the five shared genres, namely spoken,

fiction, magazine, newspaper, and academic, were included in this study for a comparable comparison.

In light of the investigation, the degree of formality, collocational patterns, and grammatical patterns were used as the criteria to distinguish the three near-synonyms by analyzing the data that emerged from the corpora. By way of illustration, the degree of formality was determined based on the frequencies of each synonym occurring across different genres. In addition, the MI test was used as the corpus statistics to identify the strength of collocations as the co-occurrence of words is statistically significant if the MI score is at least 3 (Hunston, 2002). Only the first 10 noun collocates that frequently co-occur with the target synonyms in each corpus source were examined. Since the target synonyms were adjectives, looking at how these words modified nouns might yield some useful patterns for EFL students to improve their writing and translation skills, and provide insights to investigate the connotations of each synonym. Therefore, only noun collocates were analyzed in this study. To investigate the pattern of ADJECTIVE + NOUN, one right-span of the node was set as the criterion to search for noun collocates of the target synonyms. Moreover, the concordance lines extracted from both COCA and BNC of each synonym were taken into consideration in order to enhance the understanding of their usage and meanings in particular contexts. Lastly, in terms of the grammatical patterns, only prepositional patterns of each synonymous adjective were investigated. Subsequently, the researcher re-examined the findings to increase the reliability of interpretation.

4. Results and Discussion

The synonymous adjectives *naughty*, *rebellious*, and *disobedient* are analyzed and presented according to their lexical information, that is, formality and collocational patterns, and their syntactic information; that is, grammatical patterns. The detailed analysis with rationales is as follows.

4.1 Degrees of Formality

In terms of formality, the distributions across five different genres in BNC, namely spoken, fiction, magazine, newspaper, and academic were explored to represent British English. The results were ranked in terms of its frequency per million which is a normalized frequency.

Table 2. Comparison of frequency of *naughty*, *rebellious*, and *disobedient* in different text types in BNC

Genre	<i>Naughty</i>		Genre	<i>Disobedient</i>		Genre	<i>Rebellious</i>	
	Freq.	Per million		Freq.	Per million		Freq.	Per million
SPOKEN	231	23.18	FICTION	19	1.19	MAGAZINE	28	3.86
FICTION	186	11.69	ACADEMIC	13	0.85	ACADEMIC	37	2.41
MAGAZINE	44	6.06	SPOKEN	3	0.30	NEWSPAPER	22	2.10
NEWSPAPER	45	4.30	MAGAZINE	2	0.28	FICTION	33	2.07
ACADEMIC	33	2.15	NEWSPAPER	0	0.00	SPOKEN	6	0.60
Total	539			37			126	

Table 2 illustrates that of these three synonyms, *naughty* occurs with the highest frequency in BNC (539 tokens) which is higher than *rebellious* (126 tokens), and *disobedient* (37 tokens), respectively.

According to Table 2, it can be assumed that many non-academic disciplines prefer *naughty* to *disobedient* and *rebellious* because of the high rate of occurrence of *naughty* in spoken (231 tokens) and fictions (186 tokens), and its lowest frequency in academic. To elaborate, this corpus data can support the assumption that *naughty* has a higher degree of informality reflecting its usage in colloquial and everyday language than the other target adjectives. Similarly, *disobedient* is discovered to be frequently used in a non-academic context because of its occurrence in fiction (19 tokens). However, *disobedient* can also be commonly used in the academic genre (13 tokens) signalling its usage in a formal context. More importantly, *rebellious* tends to be used in formal contexts because of its frequent occurrence in magazines (28 tokens) and academic (37 tokens).

With this assumption, these synonyms are considered near-synonyms because they cannot replace each other in some registers in British English. In order to strengthen the reliability of the study, the frequency of the occurrences of the target synonyms across five different genres in COCA, namely fiction, magazines, spoken, newspapers, and academic, are compared to validate the degree of formality of these three synonyms in American English.

Table 3. Comparison of frequency of *naughty*, *rebellious*, and *disobedient* in different text types in COCA

Genre	<i>Naughty</i>		Genre	<i>Disobedient</i>		Genre	<i>Rebellious</i>	
	Freq.	Per million		Freq.	Per million		Freq.	Per million
FICTION	460	3.89	FICTION	86	0.73	ACADEMIC	383	3.20
MAGAZINE	459	3.64	ACADEMIC	75	0.63	MAGAZINE	386	3.06
SPOKEN	288	2.28	MAGAZINE	44	0.35	FICTION	320	2.70
NEWSPAPER	228	1.87	SPOKEN	20	0.16	NEWSPAPER	309	2.54
ACADEMIC	59	0.49	NEWSPAPER	16	0.13	SPOKEN	227	1.80
Total	1,494			241			1,625	

Table 3 shows that among the three synonyms, *rebellious* occurs with the highest frequency in COCA (1,625 tokens) which is higher than *naughty* (1,494 tokens), and *disobedient* (241 tokens), respectively. Interestingly, this sequence is not aligned with those found in BNC.

As illustrated in Table 3, *naughty* is found most frequently in fiction (460 tokens), followed by magazines (459 tokens). The result indicates that *naughty* is commonly used in an informal context, reflecting its usage in fictional and conversational language. In a similar vein, the occurrence of *naughty* in American academic contexts is the lowest (59 tokens), which is consistent with the findings found in BNC. Thus, the informality of *naughty* is confirmed in both British and American English. However, *naughty* in COCA does not appear as high-frequently in spoken English as it does in BNC. Moreover, there is a clear trend that *disobedient* occurs with the greatest frequency in fiction (86 tokens), followed by academic (75 tokens). The result reveals that *disobedient* is commonly used in informal English which is in line with the occurrence of *disobedient* in BNC. More importantly, it is apparent that *rebellious* has the highest frequency occurring in academics (383 tokens), followed by magazines (386 tokens). The high rate of occurrence of *rebellious* in the academic genre and its low frequency in spoken can support the assumption that it has a higher degree of formality than the other

target adjectives. This denotes that the usage of *rebellious* in American formal contexts is common, which is corresponding to the findings in BNC. Interestingly, it was observed that the occurrences of both *disobedient* and *rebellious* in spoken American English is low, so it can be assumed that they were not commonly used in communication based on their frequencies in the spoken genres.

Thus, the degree of formality of the three synonyms in both COCA and BNC suggests that these synonyms differ in terms of the style or word register of the contexts in which they occur, confirming their status as near-synonyms. The findings of this study complied with previous research (Biber et al., 1998; Cai, 2012; Jackson & Amvela, 2000; Moon, 2010; Phoocharoensil, 2020; Shaw, 2011) that a corpus tool is considered as a useful tool in analyzing the distributions across genres of the synonyms in order to distinguish them effectively based on formality.

4.2 Collocations of *Naughty*, *Rebellious*, and *Disobedient*

The second part will focus on the collocational differences of these synonyms. In this study, MI scores in the following tables show that the target synonyms are significantly collocated with their collocates since an MI score is higher than 3 (Hunston, 2002).

Table 4. Noun collocates of the synonyms *naughty*, *disobedient*, and *rebellious* in BNC

Rank	<i>Naughty</i>			<i>Disobedient</i>			<i>Rebellious</i>		
	Noun collocate	Frequency	MI value	Noun collocate	Frequency	MI value	Noun collocate	Frequency	MI value
1	boy	51	boy 41 8.48	child	3	child 2 7.08	attitude	4	attitude 3 7.91
			boys 10 7.16						children 1 5.13
2	child	21	child 11 5.68	schoolboy	1	11.65	streak	3	11.70
			children 10 4.59						
3	girl	20	girl 18 7.10	stomach	1	9.13	officer	3	7.31
			girls 2 4.61						
4	word	10	word 4 4.58	tools	1	9.02	peasant	3	peasant 1 8.20
			words 6 4.81						
5	schoolboy	7	10.59	youth	1	8.25	determination	2	8.43
6	schoolgirl	6	schoolgirl 4 10.54	boys	1	7.70	nature	2	5.73
			schoolgirls 2 11.30						
7	things	6	4.02	hair	1	6.89	students	2	6.02
8	bits	5	7.41	mean	1	5.22	youth	2	7.48
9	dog	4	5.81	members	1	5.78	teenager	2	teenager 1 9.15
									teenagers 1 9.00
10	daddy	3	7.29	-	-	-	Yorkshire men	1	14.31

According to Table 4, it is clearly shown that *naughty* in British English tends to co-occur with the noun collocates that are associated with young children such as *child(ren)*, *boy(s)*, *girl(s)*, *schoolboy*, and *schoolgirl(s)*, supporting the definition provided in the dictionary. To be more specific, *naughty* has the highest tendency to co-occur with *boy(s)* (51

tokens), followed by *child* (21 tokens), and *girl(s)* (20 tokens), respectively. Semantically, it is interesting to point out that boys are evidently perceived as naughtier than girls, reflecting the gender stereotype. To enhance the understanding of the collocation behaviors in contexts and consequently decrease the thesaurus errors (Ferris, 2009), some examples of the concordance lines from BNC are presented.

(1) Ta da, don't be a **naughty boy** again. I open the second letter, would you please appear in court [*spoken*]

(2) the Colonel's lady whined to her husband.' Was I a **naughty girl** to ask for champagne?' [*fiction*]

(3) If they thought he was going to stand in front of them like a **naughty schoolboy**, they were mistaken.' [*fiction*]

As shown in (2), it can be assumed that *girl* in the corpus data does not only mean young children, but also adult women as *naughty* in (2) is used humorously to describe an adult's bad or socially unacceptable action. Thus, this assumption supports the definition in the dictionary. Moreover, it is noticeable that *naughty* can co-exist with non-living phenomena such as *word(s)*, *things*, and *bits*. Significantly, the findings reveal that *naughty* can be used to modify inanimate objects to convey sexual feelings and thoughts, or behaviors and languages that are not socially acceptable, which confirms the definition provided in the dictionary as illustrated in the following concordance lines from BNC.

(4) "What's that then?" **Naughty words**, Mum. Not very lady-like at all. What would Davey think? [*fiction*]

(5) is a symptom of your sexual frustration. You want your nice boyfriend to do **naughty things** -- the things this girl does in your fantasies. [*newspaper*]

(6) men looking too excited. # Naughty Magazines are allowed to show a man's **naughty bits** only if they resemble a map of the Mull of Kintyre, Scots-born Pauline [*newspaper*]

Likewise, Table 4 highlights the fact that *disobedient* shares some of the frequent collocates with *naughty* as it tends to co-occur with *child(ren)*, *schoolboy*, and *boys* as well. Similar to *naughty*, *disobedient* frequently co-exists with non-human noun collocates such as *stomach*, *tools*, and *hair*. However, *disobedient* does not denote sexual connotations. To elaborate, *disobedient* is used to personify the objects to signal the act of refusal against the users' intentions, as shown in concordances (7) and (8).

(7) In here they can't get me. And I'm safe from those **disobedient tools** too. They've had their day." [*fiction*]

(8) could sense the various cuts and nicks on my queasy jawline, my short but **disobedient hair**, the unhappy hang of my uniform, my lustreless black boots. [*fiction*]

Crucially, this usage of *disobedient* is not commonly exemplified in the dictionary despite the fact that it frequently co-occurred with non-human; therefore, it can be suggested that corpus tools can offer a new research perspective to explore language use beyond native-speaker intuition (Hunston, 2002).

Considering the noun collocates of *rebellious*, they are apparently different from *naughty* and *disobedient* because *rebellious* frequently co-exists with the words that reveal a greater sense of seriousness than the other two synonyms, for example, *streak*, *officer*, *peasants*, and *determination*, as exemplified in (9), (10), and (11).

(9) Young people with curiosity and a ***rebellious*** streak are only too ready to experiment. [*newspaper*]

(10) This ugly crime against Lebanon and its legitimacy followed a series of threats by the ***rebellious*** officer [*newspaper*]

(11) In 1549 the citizens of Exeter believed the ***rebellious*** peasants to be intent on pillage; Norwich merchants concealed their valuables and fled. [*academic*]

In addition, confirming the definitions provided in the dictionary, *rebellious* is likely to collocate with adults or young adults, for example, *officer*, *Yorkshiremen*, *peasants*, *youth*, and *teenager(s)*, while *naughty* and *disobedient* are more frequently collocated with younger groups of people. In terms of semantic preference, *rebellious* tends to create a sense of suppression and disagreement with people in authority when it co-exists with its noun collocates indicating opposed groups of people, as evidenced in the concordances (12) and (13).

(12) the third youngest president of all time and five years younger than the voice of ***rebellious*** youth, Bob Dylan. [*newspaper*]

(13) 1489 the earl of Northumberland lost his life at the hands of a force of ***rebellious*** Yorkshiremen when he was trying to collect the subsidy granted that year [*academic*]

As can be seen in the concordance lines, the corpus data can reveal some historical contexts and events that happened in a particular setting. It is also observed that *rebellious* can co-occur with abstract nouns, for instance, *attitude*, *determination*, and *nature*, to indicate a strong will of the opposition, as illustrated in (13) and (14).

(13) But maybe that's what the pop world needs -- a bit of ***rebellious*** attitude. 'Disagreed' is just having a pop at the governments of [*magazine*]

(14) And from somewhere inside she felt a spurt of ***rebellious*** determination. 'I'll make it work. I will...' And then [*fiction*]

Even though the target synonyms may share core meaning, they are different in noun collocates to some extent which could yield different shades of meanings in some contexts. Therefore, regarding the different denotative meanings of *naughty*, *disobedient*, and *rebellious*, it can be assumed that they are loose synonyms since they cannot substitute each other in all possible contexts.

In order to compare the collocations of these three synonymous adjectives across the varieties of English, the data obtained from COCA representing American English were analyzed as listed in Table 5.

Table 5. Noun collocates of the synonyms *naughty*, *disobedient*, and *rebellious* in COCA

Rank	<i>Naughty</i>			<i>Disobedient</i>			<i>Rebellious</i>					
	Noun collocate	Frequency	MI value	Noun collocate	Frequency	MI value	Noun collocate	Frequency	MI value			
1	boy	73	boy 59	7.11	child	18	child 8	6.39	teenager	41	teenager 29	9.76
			boys 14	5.79			children 10	5.84			teenagers 6	7.42
											teen-ager 6	10.98
2	child	54	child 30	5.66	son	8	6.75	spirit	28	spirit 25	7.18	
			children 24	4.47						spirits 3	6.28	
3	girl	37	girl 31	6.00	kid	4	kid 1	4.53	youth	21	7.31	
			girls 6	4.35			kids 3	4.96				
4	word	37	word 20	5.42	daughter	4	daughter 2	5.40	streak	19	9.43	
			words 17	5.03			daughters 2	7.82				
5	dog	32	7.05	wife	3	5.38	slave	17	slave 4	6.48		
									slaves 13	8.33		
6	things	29	4.23	dog	3	6.27	son	15	4.90			
7	bits	26	9.41	organ	2	9.02	republics	14	10.19			
8	list	17	5.63	legs	2	6.53	nature	14	5.33			
9	corner	8	5.74	hair	2	5.19	behavior	14	behavior 11	5.31		
									behaviors 3	5.40		
10	thoughts	7	5.65	act	2	4.92	teen	13	8.26			

As can be seen in Table 5, the collocations obtained from COCA have more varieties than those drawn from BNC. It is apparent that the first most frequent noun collocates of *naughty* and *disobedient* in COCA are similar to those found in BNC. Notably, *boy(s)* is the most frequent noun collocate of *naughty*, and *child(ren)* is the most frequent noun collocate of *disobedient*. On the other hand, the first most frequent noun collocate of *rebellious* in COCA, that is, *teenager(s)*, is different from BNC (i.e., *attitude(s)*).

In a similar vein, *naughty* is frequently used with young children (e.g., *boy(s)* and *girl(s)*), and, similarly, boys are also stereotypically viewed as the naughtier gender due to the higher tendency of *boy(s)* than *girl(s)* in COCA. Hence, these findings are consistent with the data found in BNC. Moreover, the collocation *naughty corner*, supports the assumption that *naughty* tends to be involved with children as *naughty corner* is used as a common form of punishment to discipline children.

(15) 'What Would the Supernanny Do?' Supernanny would send him to his *naughty corner* and say, 'That is unacceptable.' [*fiction*]

Interestingly, it is noteworthy to point out that *naughty* in the American context can be associated with Christmas because it has a high tendency to co-exist with *list*, suggesting the common usage of *naughty* in the festive event as illustrated in (16).

(16) Hi. I got the stomach flu on Christmas Day because I was on the ***naughty list*** this year. [*spoken*]

Furthermore, it is noticed that *naughty* is also used to signify a sexual connotation in the American context as it tends to collocate with *word(s)*, *thing(s)*, *bit(s)*, and *thoughts*, as exemplified in (17) and (18).

(17) "Whisper some sexy compliment or all the ***naughty things*** you're going to do to him when you get home, or lean [*magazine*]

(18) the relationship platonic, she had to admit that she'd been having some very ***naughty thoughts*** about her hot new boss. [*fiction*]

In terms of *disobedient*, the noun collocates found in COCA are relatively similar to BNC. The occurrences of *disobedient* in American English is often related to dominance-submission as it is frequently collocated with *kid(s)*, *child*, *son*, *daughter*, and *wife*. As illustrated in (19) and (20), it can be inferred that these collocates mirror patriarchy and dominance-submission in families.

(19) but he has left his mark, for in pressing the blood-stained key to his ***disobedient wife's*** forehead, he has left there the tell-tale sign of her desire branded [*academic*]

(20) the hand and forcibly put him to bed. If he wanted to be a ***disobedient child***, I was forced to play the role of the strict parent. [*fiction*]

Similar to BNC, it is observed that *disobedient* in COCA can be commonly used with noun collocates associated with body organs such as *organ*, *legs*, and *hair* to give them the human characteristic of behaving against the owner's will, as shown in the following concordances.

(21) "Just forget it, " he said, as much to his ***disobedient organ*** as to his unintentionally sexy consultant. "It's a cute idea [*fiction*]

(22) I walked, or more likely wobbled, out of Sunbrook Retirement Residence, my ***disobedient legs*** carried me back to the park and the weathered redwood bench [*fiction*]

Notably, *rebellious* in COCA shares some common noun collocates with BNC, for example, *teenager(s)*, *youth*, *streak*, and *nature*. Despite the shared noun collocates, it is observed that *rebellious* has the highest tendency to co-occur with *teenager(s)* (41 tokens) which appeared to occur with low frequency in BNC. More importantly, it can be assumed that the occurrence of *rebellious* with *teenager(s)*, *youth*, and *teens* manifests the higher usage of *rebellious* with young adults than the other two synonyms, as evidenced in (23).

(23) my vacation to work for free helping strangers," recalls Nathan, a typically **rebellious** teenager who usually avoided spending time with his parents and sisters. [newspaper]

Furthermore, the data obtained from COCA reflect some parts of the American history regarding African American slavery as *rebellious* is discovered to co-occur with *slave(s)*, signifying the usage of *rebellious* with a higher degree of seriousness than the other two synonyms of racial discrimination and enslavement, as illustrated in (24) and (25).

(24) # 1839 # Amistad Mutiny A Spanish ship commandeered by 53 **rebellious** slaves lands on Long Island. [magazine]

(25) Two of the more notorious Americans represented are the **rebellious** slave Nat Turner and the fire-and-brimstone preacher Jonathan Edwards. [academic]

Thus, concordance analysis and the study of collocational patterns can lead to more insightful information and understanding of meaning in context because vocabulary cannot be acquired in isolation and the co-selection of words also needs to be taken into account (Cheng, 2012; Sinclair, 2004; Thornbury, 2002). It can be concluded that the collocational patterns of *naughty*, *disobedient*, and *rebellious* obtained from COCA and BNC confirms their status as near-synonyms since they could not share the same shades of meanings and cannot be used interchangeably in all contexts. Moreover, it was found that there is a consistency of the collocational patterns and connotations of the three synonyms across different varieties of English. The current findings are generally in agreement with those in previous studies (Aroonmanakun, 2015; Biber et al., 1998; Chung, 2011; Jackson & Amvela, 2000; Moon, 2010; Sumintana & Tawilapakul, 2019).

4.3 Grammatical Patterns of *Naughty*, *Rebellious*, and *Disobedient*

The last important criterion used to distinguish the synonyms is grammatical patterns. Due to the fact that *naughty*, *rebellious*, and *disobedient* are adjectives, they tend to occur as a modifier in attributive positions, postpositive positions, and predicative positions to modify nouns. Therefore, the target grammatical pattern in the study is ADJECTIVE + PREPOSITION. The possible grammatical patterns of *naughty*, *rebellious*, and *disobedient* were elicited from Word Sketch function of SkE in BNC corpus, as shown in Table 6.

Table 6. The grammatical patterns of *naughty*, *rebellious*, and *disobedient* in BNC

Synonyms	Adjective + Preposition	Examples from concordance lines (BNC)
<i>naughty</i>	<i>naughty</i> by	Are these just bad vibes and sour grapes or is hip hop just too naughty by nature for the mainstream?
	<i>naughty</i> of	It's naughty of your mother not to let you have them.
	<i>naughty</i> with	So prodigal and naughty with caresses ,
<i>rebellious</i>	<i>rebellious</i> in	people who have experienced unemployment are more rebellious in spirit does not itself explain why this occurs.
	<i>rebellious</i> about	Tories who feel rebellious about the railways,
	<i>rebellious</i> towards	He's a very like er rebellious towards rock and roll.

	<i>rebellious</i> against	his father or some other relation was among the most bitterly <i>rebellious <u>against</u> a system</i> that could tolerate such things.
<i>disobedient</i>	<i>disobedient</i> to	were the greatest enemies to monarchy and most <i>disobedient <u>to</u> Kings</i> ,

Regarding their grammatical structures in the forms of ADJECTIVE + PREPOSITION, as shown in Table 6, it can be observed that the three synonyms are likely to be followed by different prepositions. Interestingly, the preposition *against* that co-occurs with *rebellious* (e.g., *rebellious against a system*) could reflect the distinctive register of this word as it could enhance the sense of conflict and opposition, which is corresponding to the noun collocates of *rebellious*. Thus, *naughty*, *disobedient*, and *rebellious* are loose synonyms since they do not share the same grammatical patterns in some particular contexts. In order to cross-check the data in BNC, the ADJECTIVE + PREPOSITION patterns of the three synonyms in COCA were investigated.

Table 7. The grammatical patterns of *naughty*, *rebellious*, and *disobedient* in COCA

Synonyms	Adjective + Preposition	Examples from concordance lines (COCA)
<i>naughty</i>	<i>naughty</i> by	Women are nice and men are <i>naughty <u>by</u> nature</i> .
	<i>naughty</i> with	We have to do a bit <i>naughty <u>with</u> the red vibes</i> .
	<i>naughty</i> to	Kissing was very <i>naughty <u>to</u> the Japanese</i> . Santa thinks it's <i>naughty <u>to</u> have wine</i> before nine.
<i>rebellious</i>	<i>rebellious</i> with	He's going to get a little more <i>rebellious <u>with</u> his parents</i>
	<i>rebellious</i> by	We're teens -- <i>rebellious <u>by</u> nature</i> .
	<i>rebellious</i> about	He's got something to be <i>rebellious <u>about</u></i> ... namely, being gay in a homophobic society.
	<i>rebellious</i> to	He was argumentative and <i>rebellious <u>to</u> his father</i>
	<i>rebellious</i> against	Maha and Hanan have always been <i>rebellious <u>against</u> background, tradition, and taboos</i> of all sorts.
<i>disobedient</i>	<i>disobedient</i> to	I will become <i>disobedient <u>to</u> God's call</i> for holiness in my life.

According to Table 7, despite some shared prepositions (i.e., *with* and *to*), it is apparent that the target synonyms have different grammatical patterns regarding ADJECTIVE + PREPOSITION. Moreover, it was discovered that there are some similar and different usages of prepositions that follow the three synonymous adjectives across two sources of corpora representing two varieties of English. To illustrate, Table 7 reveals some unique grammatical patterns in COCA that do not appear in BNC. For instance, *naughty* in COCA can be followed by *to* + *infinitive* to show the action (e.g., *naughty to have wine*) and *to* + *noun* to indicate a person whom a speaker behaves badly to (e.g., *naughty to the Japanese*), on the other hand, *naughty of* only exists in BNC. It is also interesting to point out that *rebellious* can be followed by a wider variety of prepositions in COCA than in BNC, namely *with*, *by*, and *to*. However, it cannot be generalized that *disobedient in* or *rebellious to* do not exist in British English because of the limitation of the corpus evidence that is not capable of collecting all of the language used in natural-occurring contexts (Aston, 2001). Results such as these align with those found in Gu (2017) and Petcharat and Phoocharoensil (2017). Since the three synonyms do not share

all the same grammatical patterns, thus they cannot replace each other in all grammatical contexts, confirming their status as near-synonyms.

5. Conclusion and Pedagogical Implication

In conclusion, this study sheds some light on how to implement corpus data into analyzing the similarities and differences of English synonyms. Based on the contrastive corpus analysis across BNC and COCA in this study, the synonymous adjectives *naughty*, *rebellious*, and *disobedient* yield interesting results to confirm that most English synonyms are loose synonyms (Palmer, 1997) because they cannot be used interchangeably in every context (Jackson & Amvela, 2000) with regard to their distinctiveness of collocations, degrees of formality, and grammatical patterns. Regarding distributions across five different genres, it can be concluded that these three synonymous adjectives cannot substitute each other in all contexts due to their different degrees of formality. *Naughty* most frequently occurs in informal contexts as it has the highest tendency to appear in non-academic contexts in both American and British English. Conversely, *disobedient* and *rebellious* most frequently occur in more academic disciplines reflecting its usage in formal contexts. Therefore, the distinction in genres and formality highlighted with the assistance of corpus tools are important and useful criteria to effectively differentiate synonyms (Szudarski, 2018). In addition to genre differences, collocation analysis by looking at the top 10 most frequent noun collocates of each synonym is considered to be another vital criterion to distinguish near-synonyms (Cheng, 2012; Sinclair, 2004; Thornbury, 2002). Based on the discovered collocational patterns of the target synonyms, they are still different in noun collocates signaling different lexical patterns and usages in naturally occurring contexts; thus, it can be concluded that *naughty*, *disobedient*, and *rebellious* cannot be used interchangeably in all contexts. Interestingly, some of the emerged corpus data reveal that *naughty* and *disobedient* tend to share some noun collocates such as *child(ren)* and *boy(s)*, which indicate overlapping meanings and support the definition in the dictionary that *naughty* and *disobedient* are commonly used with children to suggest undesired behaviors. Crucially, one of the differences between *naughty* and *disobedient* is that *naughty* tends to denote sexual connotation when it is used with inanimate objects such as *things* and *words* whereas *disobedient* is used to personify the objects to signal the act of refusal against the users' intentions such as *disobedient tools* and *disobedient legs*. It is also noteworthy to point out that this collocational pattern of *disobedient* is not widely covered or exemplified in the dictionary; thus, corpus tools can offer a new research perspective to explore language use beyond native-speaker intuition (Hunston, 2002). On the other hand, *rebellious* is more likely to be used with teenagers. Nevertheless, the corpus data drawn from both BNC and COCA suggest the usage of *rebellious* with a higher degree of seriousness than the other two synonyms when it co-occurs with *slave(s)*, *streak*, and *determination*. Lastly, according to the grammatical pattern ADJECTIVE + PREPOSITION elicited from the corpora, the three synonyms are likely to be followed by different prepositions which confirm their status as near-synonyms.

To some extent, these findings can have a profound impact in the fields of English language teaching, especially in vocabulary teaching, as it is apparent that corpora serve as a useful tool to explore the authentic language patterns and help enhance learners' metalinguistic awareness through investigating concordance lines and collocation patterns (Shaw, 2011). By following the above corpus-based approach, ELT practitioners and L2 learners will be able to distinguish synonyms effectively, apart from solely consulting dictionaries or native-speaker

intuition. Moreover, exploring language patterns through corpus data can enhance learners' lexical knowledge and lexical competence (Liu, 2010; Phoocharoensil, 2010; Shahzadi et al., 2019). In light of classroom implications, providing learners with concordance lines and collocational patterns extracted from COCA and BNC when introducing new vocabulary to them is considered to be more useful than providing only a single word (Jirananthiporn, 2018) as they can exemplify the real-life language patterns. It is also noteworthy to point out that learning vocabulary through the concordances and collocational patterns can broaden the learners' horizons as they may incidentally learn cultures, beliefs, ideologies, and history of the target language, especially when two corpora across English varieties are compared. For example, the usage of *naughty list* in the American context signals Christmas and the pattern of *rebellious slave(s)* signifies African American slavery. Moreover, frequency lists showing degrees of formality of each word and grammatical patterns should be provided to learners in order to help them overcome difficulties in distinguishing near-synonyms, apply the synonyms in the correct contexts, reduce the thesaurus errors, and consequently, improve their writing and translation skills (Ferris, 2009; Phoocharoensil, 2010; Sasum & Weeks, 2018). To elaborate, it would be more beneficial for learners to be exposed to the ADJECTIVE + PREPOSITION patterns such as *rebellious against* and *naughty to* so that they will be aware of the correct usage of each synonym and use corpora as a useful resource to refine their writing. In conclusion, the incorporation of corpus evidence in EFL/ESL material development should be advocated as corpora can provide some additional data that is non-existent in the dictionary or beyond native-speaker intuition (Aroonmanakun, 2015; Sinclair, 1991; Szudarski, 2018).

The Author

Punyapa Boontam Lecturer, Division of Foreign Languages, Nakhon Sawan Rajabhat University, Nakhon Sawan, Thailand. Currently undertaking PhD studies in English Language Teaching, Language Institute, Thammasat University. M.A. (English Language Teaching), Language Institute, Thammasat University. B.A. (English), 1st Class honors, Chiang Mai University. Areas of interest: Corpus Linguistics, Data-driven Learning, English Language Teaching.

References

- Alquahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education*, 3(3), 21-34.
- Aroonmanakun, V. (2015). Quick or fast: A corpus based study of English synonyms. *LEARN Journal: Language Education and Acquisition Research Network Journal*, 8(1), 53-62.
- Aston, G. (2001). *Learning with corpora*. Athelstan.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

- Cai, J. (2012). *Is it “great” enough? A corpus-based study of “great” and its near-synonyms*. [Master’s thesis, Ball State University].
- Cambridge University Press. (n.d.). Disobedient. In *Cambridge dictionary*. Retrieved October 10, 2020 from <https://dictionary.cambridge.org/>
- Cambridge University Press. (n.d.). Naughty. In *Cambridge dictionary*. Retrieved October 10, 2020 from <https://dictionary.cambridge.org/>
- Cambridge University Press. (n.d.). Rebellious. In *Cambridge dictionary*. Retrieved October 10, 2020 from <https://dictionary.cambridge.org/>
- Caro, K., & Mendinueta, N. R. (2017). Lexis, lexical competence and lexical knowledge: A review. *Journal of Language Teaching and Research*, 8(2), 205-213.
- Cheng, W. (2012). *Exploring corpus linguistics: Language in action*. Routledge.
- Chung, S. (2011). A corpus-based analysis of “Create” and “Produce”. *Chang Gung Journal of Humanities and Social Sciences*, 4(2), 399-425.
- Ferris, D. (2009). *Teaching college writing to diverse student populations*. University of Michigan Press.
- Francis, G., Hunston, S., & Manning, E. (1996). *Collins cobuild grammar patterns 1: Verbs*. Harper Collins.
- Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, 35, 305-327. <https://doi.org/10.1093/applin/amt015>.
- Gu, B. (2017). Corpus-based study of two synonyms-obtain and gain. *Sino-US English Teaching*, 14(8), 511-522.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
- Jackson, H., & Amvela, E. (2000). *Words, meaning and vocabulary: Meaning relations*. Athenaum Press.
- Jackson, H., & Amvela, E. (2007). *Words, meaning and vocabulary: An introduction to modern English lexicology*. Continuum.
- Jarunwaraphan, B., & Mallikamas, P. (2020). A corpus-based study of English synonyms: Chance and opportunity. *rEFlections*, 27(2), 218-245.
- Jirananthiporn, S. (2018). Is this problem giving you trouble? A corpus-based examination of the differences between the nouns ‘problem’ and ‘trouble’. *Thoughts*, 2, 1-25.
- Jones, C. & Waller, D. (2015). *Corpus linguistics for grammar: A guide for research*. Routledge.

- Kirkpatrick, A. (2007). *World Englishes*. Cambridge University Press.
- Lindquist, H. (2009). *Corpus linguistics and the description of English*. Edinburgh University Press.
- Ling, L. Y. (2005). Teaching vocabulary learning strategies: Awareness, beliefs, and practices. *A Survey of Taiwanese EFL Senior High School Teachers*. [Unpublished master's thesis, The University of Essex].
- Liu, D. (2010). Is it a chief, main, major, primary, or principal concern? A corpus-based behavioral profile study of the near-synonyms. *International of Corpus Linguistics*, 15(1), 56-87.
- Moon, R. (2010). What can a corpus tell us about lexis?. In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics*. (pp. 345-358). Routledge.
- Nation, I. S. P. (1994). *New ways in teaching vocabulary*. TESOL.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nation, I. S. P. (2017a). The BNC/COCA Level 1 word family lists (Version 1.0.0) [Data file]. <http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>
- Nation, I.S.P. (2017b). The BNC/COCA Level 8 word family lists (Version 1.0.0) [Data file]. <http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>
- O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.
- Palmer, F. R. (1997). *Semantics*. Cambridge University Press.
- Petcharat, N., & Phoocharoensil, S. (2017). A corpus-based study of English synonyms: Appropriate, proper, and suitable. *LEARN Journal: Language Education and Acquisition Research Network Journal*, 10(2), 10-24.
- Phoocharoensil, S. (2010). A corpus-based study of English synonyms. *International Journal of Arts and Sciences*, 3(10), 227-245.
- Phoocharoensil, S. (2020). Collocational patterns of the near-synonyms *error*, *fault*, and *mistake*. *The International Journal of Communication and Linguistic Studies*, 19(1), 1-17. <https://doi.org/10.18848/2327-7882/CGP/v19i01/1-17>.
- Sasum, S., & Weeks, B. (2018, May 4). *Why some Thai students cannot speak English fluently?*. RSUSSH 2018: Integrated Innovation and Startup for the Future, 361-367, Bangkok, Thailand.
- Shahzadi, A., Asghar, A., & Javed, S. (2019). Effectiveness of corpus in teaching English synonyms. *Corporum: Journal of Corpus Linguistics-CJCL*, 2(1), 51-65.

- Shaw, E. M. (2011). *Teaching vocabulary through data-driven learning*. Brigham Young University.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- Sinclair, J. (2004). *Trust the text*. Routledge.
- Sotoudehnama, E., & Soleimanifard, F. (2013). The effects of teaching vocabulary through synonymous, semantically unrelated, and hyponym sets on EFL learners' retention. *Issues in Language Teaching (ILT)*, 2(2), 27-56.
- Sumintana, N., & Tawilapakul, U. (2019, July). *A corpus-based study of the English synonyms cure and heal*. The 8th LITU International Graduate Conference: The Journey to Publication, 148-154, Bangkok, Thailand.
- Szudarski, P. (2018). *Corpus linguistics for vocabulary: A guide for research*. Routledge.
- Thornbury, S. (2002). *How to teach vocabulary*. Longman.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. John Benjamins Publishing.