



A Corpus-based Investigation into English Synonymous Verbs ‘Restrict’ and ‘Constrain’

Nateethorn Narkprom

nateethorn.nar@pcru.ac.th, Faculty of Education, Phetchabun Rajabhat University, Thailand

APA Citation:

Narkprom, N. (2024). A Corpus-based investigation into English synonymous verbs ‘restrict’ and ‘constrain’. *LEARN Journal: Language Education and Acquisition Research Network*, 17(1), 688-714.

Received
17/10/2023

Received in
revised form
28/11/2023

Accepted
09/12/2023

ABSTRACT

Through consultations with the online version of the Longman Dictionary of Contemporary English (LDOCE) and the Corpus of Contemporary American English (COCA), this study focuses on distinguishing between the two synonymous verbs ‘restrict’ and ‘constrain’, both part of Coxhead’s (2000) Academic Word List, in terms of formality and collocations. The COCA-informed findings revealed that both near-synonyms are most prevalent in academic genres, sharing an equally high level of formality, and exhibit six overlapping groups of noun collocates. Interestingly, the target synonyms, along with their left- and right-sided noun collocates, form context-specific clusters. The verb ‘restrict’ is associated with terms related to governance and rights (i.e., *laws and regulations restrict liberties*) while ‘constrain’ with governance and commerce (i.e., *rules constrain spending*). The two are also characterized by unique patterns of collocating with nouns that have opposite morphological meanings: ‘restrict’ pairs with noun collocates containing morphemes representing inbound activities (e.g., *acc-* in *access*, as in *restrict access*), while ‘constrain’ with those indicating outbound directions (e.g., *ex-* in *expansion*, as in *constrain expansion*).

Keywords: synonym, verb, collocation, genre, COCA

Introduction

Learning words is the essence of learning a new language. According to Folse (2011), vocabulary is of paramount importance in learning a foreign language. In the context of English as a Foreign Language (hereinafter, EFL), a number of English Language Teaching (hereinafter, ELT) practitioners have been led to believe that knowing as many words as possible is beneficial for students long-term (Nation, 2001). As a result, teaching synonyms has become a common practice, aimed at enhancing students' lexical repertoire. However, by prioritizing quantity over quality, teaching synonyms may become counterproductive. Laufer (1990) highlighted an interesting reason why teaching English synonyms could be problematic—average students tend to have a misconception that synonyms can be used interchangeably, or unconditionally substitutable with their equivalents, irrespective of contexts, a notion that has not been proven to be the case.

For instance, the adjective *strong* is synonymous with *powerful*. The former can modify the noun *tea*, as in *strong tea*, but not *engine*, as in **strong engine*, while the latter is vice versa—it can precede *engine*, as in *powerful engine*, but not *tea*, as in **powerful tea*. The counterproductivity of teaching synonyms does not only occur with students of low and intermediate proficiency but also with advanced learners. Martin (1984) discovered that advanced learners, who were given vocabulary instruction through synonyms, also made incorrect lexical choices, as a result of the lack of 1) stylistic, 2) syntactic, 3) collocational, and 4) semantic information around those synonyms.

In recent years, studies on English near-synonyms using a corpus-based approach have been steadily on an upward trajectory in the landscape of applied linguistics, with attempts to individualize English near-synonyms based on levels of formality, connotations, collocations, semantic prosody, and colligations (Phoocharoensil, 2021a). In other words, the corpus-based approach to English near-synonyms can provide insightful information on style, syntax, collocation, and semantics of the target synonyms, effectively addressing all of the four problems in teaching vocabulary via synonyms, as previously raised by Martin (1984).

A comprehensive review of existing literature in international peer-reviewed journals focusing on pedagogical studies of English near-synonyms indicates that past researchers typically analyzed pairs or sets of synonymous words through corpus data across three key grammatical functions: *nouns* (Jarunwaraphan & Mallikamas, 2020; Jirananthiporn, 2018; Lertcharoenwanich & Phoocharoensil, 2022; Phoocharoensil, 2020a, 2020b; Sumonsriworakun, 2022; Supanfai, 2022), *adjectives* (Ajmal et al., 2022; Aroonmanakun, 2015; Aroonmanakun & Aroonmanakun, 2023; Chaokongjakra, 2023; Imsa-ard & Phoocharoensil, 2022; Liu, 2010; Petcharat

& Phoocharoensil, 2017; Phoocharoensil, 2022; Sangeekaew, 2023), and *verbs* (Boontam & Phoocharoensil, 2022; Chaengchenkit, 2023; Imsa-ard, 2021; Kruawong & Phoocharoensil, 2022; Phoocharoensil, 2021a, 2021b; Phoocharoensil & Kanokpermpoon, 2021; Song, 2021; Sridhanyarat & Phoocharoensil, 2023; Yang, 2016).

Among the three grammatical functions, synonymous verbs, particularly high-frequency ones, have received the most attention from previous scholars. However, there have been too few research attempts to discriminate between the profiles of synonymous verbs within Coxhead's (2000) Academic Word List (hereinafter, AWL). To the researcher's best knowledge, two notable exceptions are the corpus-based discrimination of synonymous verbs, *concentrate* and *focus*, carried out by Imsa-ard (2021) and that of *evaluate* and *assess*, conducted by Sridhanyarat and Phoocharoensil (2023).

Since academic vocabulary is crucial for tertiary students' understanding and writing of academic language and subject-specific content, (Choo et al., 2017; Zwiers, 2008), an effort to provide further insights into the distinctiveness of synonymous AWL verbs through a corpus-based approach could be pedagogically rewarding. According to Nation (2001), the words in Coxhead's (2000) AWL are regarded as specialized vocabulary, not commonly found in non-technical language, but rather omnipresent across a wide range of academic genres. Consequently, several commercial English language dictionaries do not provide comprehensive information and distinguishable traits regarding synonymous AWL verbs. This shortage could be considered as a missed opportunity for learners of English, who aspire to master the usages of AWL synonymous verbs in their scholarly compositions but may be unaware of subtle differences between them (e.g., collocational patterns, semantic preferences).

Both of the synonymous verbs 'restrict' and 'constrain' were selected for this corpus-based analysis because they belong to the AWL sublists 2 and 3 (Coxhead, 2000), respectively, ranking them among the top-tier academic words in descending order from sublists 1 to 10. Furthermore, 'restrict' and 'constrain' are classified as CEFR B2 and C1 vocabulary in the Oxford 5000 core word list, representing upper-intermediate and proficient levels of usage, respectively. Hence, mastering these two synonymous verbs, with assistance of corpus-based findings, would empower either upper-intermediate or advanced learners of English to communicate more effectively in both academic and professional contexts.

Literature Review

In this section, the concept of synonyms, corpus linguistics and studies on English synonyms, and pedagogical criteria for discriminating near-synonyms will be discussed, respectively.

Synonyms

Synonyms broadly denote a likeness in meaning. Using Cruse's (1986) terminology, synonyms are "sets of words or lexical items that indicate a particular similarity" (p. 265). For examples, as per Crystal (2003), the words *range*, *selection*, and *choice* are synonymous with one another in the context of "What a nice collection of flowers" (p. 450), as all three nouns convey the denotative meaning of a group of items that involves decision-making. English synonyms are categorized into two groups: perfect and near synonyms. Perfect synonyms, also known as absolute or strict synonyms, are words with exactly the same meanings that can be replaced by their lexical equivalents across various contexts while preserving the original sense, logic, communicative purposes, or meaning (Edmonds & Hirst, 2002). However, perfect synonyms are atypical since they defy the ever-changing nature of language (Cruse, 1986), challenge the individualism of words (Clarke, 1992), and, ultimately, result in redundancy (Jackson & Amvela, 2007).

In contrast, near or loose synonyms are words with a relatively similar meaning but lack interchangeability (Edmonds & Hirst, 2002) or vary in connotation, style, expression, and structure (Cruse, 1986). Despite their close semantic relationship, their overlapping meanings are distinguishable (Phoocharoensil, 2020a, 2020b) and not substitutable across contexts (Edmonds & Hirst, 2002). For example, *die* and *pass away*, which both literally mean "stop living", are unlikely to be used interchangeably. The phrasal verb *pass away* is typically used with reference to human beings or domesticated pets, rather than plants, animals, or livestock. As a result, according to Edmonds and Hirst (2002), using the expression, *pass away*, in "*Many cattle passed away during the drought" might sound awkward (p. 7). Therefore, it is crucial for both non-native English learners and teachers to distinguish near-synonyms with similar denotative meanings, as their usage is multifaceted and context-dependent, often causing confusion among English as a foreign language users (Phoocharoensil, 2020a, 2022). As per Carter (2012), contextual awareness assists English language users in using synonyms correctly. The subsequent section will discuss certain criteria used for discriminating near synonyms in this study.

Corpus Linguistics and Studies on English Synonyms

Corpus linguistics has significantly influenced how researches on English synonyms are conducted since a corpus, as a textual archive of written and spoken language, can provide researchers with authentic data representative of a variety of linguistic patterns on both grammatical and lexical levels (Bennett, 2010). Methodology-wise, corpus linguistics enables researchers to investigate an authentic language (Lindquist, 2009) and measure it both quantitatively and qualitatively (Biber et al., 1999). Specifically, quantitative analysis using corpus-based approach, primarily involving frequency counts, makes the occurrences between target synonymous words or phrases across different genres distinguishable (Petcharat & Phoocharoensil, 2017). In addition, as per Moon (2010), synonymous words can be differentiated using a corpus-based approach by comparing between genres, word frequency, phrases, and collocation.

Criteria for Discriminating Synonyms

Relevant studies have used various criteria to discriminate near synonyms across different contexts (i.e., connotations, formality, and collocations). Firstly, synonymous words can be differentiated from each other based on their connotations. Room (1985) distinguished between the connotations of *woods* and *forest*, suggesting that *woods* convey a smaller and more human-centric sense than the latter, while *forest* feels larger, wilder, and more detached from human settlements than the former. Moreover, because *woods* is often in close proximity to people's habitats, it tends to connote a less biodiverse sense compared to *forest*. Secondly, style or formality can be used as a criteria to distinguish near-synonyms from each other. Using the examples from Edmonds and Hirst (2002), the loose synonyms *pissed*, *drunk*, and *inebriated* vary in terms of formality. The adjectives *pissed* and *inebriated* are considered informal and formal, while *drunk* falls somewhere in between. Recent corpus-based studies have used formality as a criterion to distinguish synonymous verbs. For instance, Phoocharoensil (2021a, 2021b) compared two pairs of synonymous verbs: *persist* versus *persevere* and *foresee* versus *predict*, across eight textual genres in COCA. His analyses revealed that each verb from those pairs exhibited a different degree of formality. The verbs *persist* and *predict* are most commonly found in academic texts, while *persevere* and *foresee* are most prevalent in webpages, which are considered to have lower formality, compared to academic genres. The third criterion often used by linguists for discriminating between near-synonyms is collocations. The term collocation is conceptualized as words that can be attached to another word, dependent on context (Baker et al., 2006). The examples of near-synonyms

that require different collocational patterns are: *task* and *job*. Specifically, the verb *face* can proceed the noun *task* in a verb-object pattern but not *job*. In other words, the expression “face a daunting task” is more suitable than “*face a daunting job”, in terms of collocations (Edmonds & Hirst, 2002, p. 111). In corpus linguistics, researchers statistically analyze collocations using measurements such as mutual information (hereinafter, MI scores), log-likelihood, or the Z-score, among others (Flowerdew, 2012; Saito, 2020). According to Murphy (2009), individual words are likely to co-occur with specific sets of collocations, implying that even near synonyms with a similar denotation can also prefer distinctive groups of collocations (Phoocharoensil, 2022). Many relevant studies have adopted collocational analysis as a criterion for individualizing synonymous verbs. For example, as previously discussed, Phoocharoensil (2021a, 2021b) not only examined the levels of formality but also the collocations of the two pairs of synonymous verbs—*persist* versus *persevere*, and *foresee* versus *predict*. In Phoocharoensil (2021a), the verb *persist* is primarily collocated with nouns representing unpleasantness (e.g., *symptom*, *rumor*, *gap*, *myth*, *drought*, *inequality*, *tension*, *poverty*, *racism*), while *persevere* tends to co-occurs with noun collocates denoting Christianity such as *God*, *saints*, *prayer*, *guard against evil*, *spiritual aridity*, and *a Catholic school*. According to Phoocharoensil (2021b), both *foresee* and *predict* share a similar group of nouns associated with adversity. Nevertheless, the number of skepticism-oriented noun collocates for *foresee* (e.g., *collapse*, *danger*, *decline*, *difficulty*, *disaster*, *harm*, *inability*, *recession*, *revolution*, *shortage*) is higher than that for *predict* (e.g., *collapse*, *demise*, *earthquake*, *mortality*). It was concluded that *foresee* is more likely to connote negativism than *predict*, which leans more toward neutrality. The most recent corpus-based studies on synonymous verbs within AWL (Imsa-ard, 2021; Sridhanyarat & Phoocharoensil, 2023) also reveal that each AWL verb has individual preference to noun collocates. For instance, in Imsa-ard’s (2021) study, even though the most common noun collocates of the AWL verbs *concentrate* and *focus* are semantically related to geography (e.g., *America*, *Iraq*, *Japan*, *Russia*), each of them is also collocated with other nouns from different semantic groups: *concentrate* with nouns denoting exertion (e.g., *effort*, *difficulty*), while *focus* with scientific experiments (e.g., *aspect*, *prevention*, *lens*, *laser*, *beam*). Similarly, in Sridhanyarat and Phoocharoensil’s (2023) study, both AWL verbs, *assess* and *evaluate*, have seven semantic groups overlapping with each other: effect (e.g., *effect*, *effectiveness*, *impact*), quality (e.g., *performance*, *progress*), harm (e.g., *risk*), extent (e.g., *extent*), change (e.g., *changes*, *difference*), ability (e.g., *ability*, *knowledge*, *skills*), and learning (e.g., *learning*). The verb *evaluate* is also associated with two other semantic groups of noun collocates, namely, human (e.g., *candidates*, *players*, *teachers*), and statement (e.g., *claims*, *evidence*, *information*). According to the findings of these corpus-based studies, synonymous verbs, despite sharing similar denotative meanings, exhibit subtle differences in connotations, formality, and collocations.

Following relevant literature and past researchers, the present study employs both the degree of formality and patterns of collocation as criteria for distinguishing between the synonymous verbs ‘restrict’ and ‘constrain’, aiming to answer the subsequent research questions.

- 1) How are the synonymous verbs ‘restrict’ and ‘constrain’ distributed across all eight sections of the COCA textual genre?
- 2) What are the most frequent noun collocates that co-occur with ‘restrict’ and ‘constrain’, either in subject-verb or verb-object patterns?

Methodology

Data Collection and Data Analysis

To discriminate between the two synonymous verbs ‘restrict’ and ‘constrain’, the researcher utilized the Corpus of Contemporary American English (hereinafter, COCA), which is, as per Phoocharoensil (2022), one of the most established and extensively consulted corpora in the field of corpus linguistics. Representative of American English, COCA is a one-billion-word text archive, evenly collected from eight genres: spoken, fiction, magazines, newspapers, academic journals, television or movie subtitles, blogs, and websites. Not only is COCA extensively used as a reliable source of data in corpus linguistics, but it also makes significant contributions in the field of ELT for three major reasons: first, through COCA, researchers can explore the collocational patterns and usages of a target word across different language genres; second, as the size of COCA increases annually with additional data, it is considered an up-to-date source for contemporary English, providing valuable information for ELT practitioners; and third, the textual data in COCA is representative of authentic language, facilitating autonomous and inductive learning for students (Phoocharoensil, 2022).

Taking the aforementioned reasons into consideration, the researcher chose to consult COCA to address the two research questions, both of which aim to discriminate between the near-synonyms ‘restrict’ and ‘constrain’. For the first research question, the two target verbs were studied in terms of their occurrences and distributions across eight text genres in COCA. As for the second question, the most common noun collocates of the two target verbs, either in subject-verb or verb-object patterns, were extracted using COCA’s built-in collocation feature. In this study, MI score, a statistical value that regulates “the chance of two words co-occurring in consideration of their frequencies in co-occurring with all the others in the corpus” (Lui & Lei, 2018, p. 6), was used to extract the top-30 noun collocates of the synonymous verbs ‘restrict’ and ‘constrain’, with a threshold of ≥ 3 , a significant value for collocational association, as established by Cheng (2012). Subsequently, the

top-30 noun collocates of the two synonymous verbs ‘restrict’ and ‘constrain’ were semantically grouped, with those conveying similar denotations placed in the same category. In the process of semantically grouping the noun collocates of the target synonyms, the researcher received assistance from a native English speaker with over ten years of experience in ELT practice to ensure the accuracy of data analysis. Furthermore, the noun collocates of the two synonymous verbs ‘restrict’ and ‘constrain’ were compared and analyzed to identify context-specific collocational clusters that emerged from the data, along with semantic preferences worth discussing.

Results and Discussion

In Table 1, the shared definitions and sentence examples of the two synonymous verbs ‘restrict’ and ‘constrain’, which are presented in the online version of Longman Dictionary of Contemporary English (2023), hereinafter referred to as LDOCE, are illustrated. The pair of target synonyms is also discussed in terms of their overall meaning, connotative differences, and grammatical usage.

Table 1

Definitions and Usage of ‘Restrict’ and ‘Constrain’ in LDOCE Online

Synonyms	Definitions	Sentence Examples
Restrict	1. [transitive] to limit or control the size, amount, or range of something	The new law restricts the sale of handguns.
	2. [transitive] to limit someone’s actions or movements	The cramped living conditions severely restricted the children’s freedom to play.
Constrain	1. [transitive] to stop someone from doing what they want to do	Financial factors should not constrain doctors from prescribing the best treatment for patients.
	2. [transitive] to limit something	Poor soil has constrained the level of crop production.

(LDOCE Online, 2023)

Based on the definitions and examples, as previously shown in Table 1, although the verbs ‘restrict’ and ‘constrain’ are near-synonyms, as both of them mutually mean “to limit something” and are used transitively, subtle connotative differences between them are noticeable. While ‘restrict’ typically signifies confining something within certain rules, as in *The new law restricts the sale of handguns*, ‘constrain’ implies a more emphatic limitation, often

influenced by external factors, as in *Poor soil has constrained crop production levels*. Furthermore, it is worth noting that the online version of LDOCE categorizes ‘restrict’ as a medium-frequency word and ‘constrain’ as a lower-frequency word, respectively. This classification is partly due to their significant presence in the AWL. According to Nation (2001), the words in Coxhead’s (2000) AWL are regarded as specialized or technical vocabularies that are not commonly found in general language but are relatively prevalent in academic texts. In other words, ‘restrict’ and ‘constrain’, emerge as significant vocabulary in Coxhead’s (2000) specialized corpus of academic texts but occur infrequently in that of LDOCE. Consequently, unlike other high-frequency words, the dictionary lacks information on the typical word combinations or collocational patterns for these two synonymous verbs. This hinders learners from relevant academic disciplines and potentially professional writers from gaining comprehensive insights and usages into the two specialized synonymous verbs when needed.

Another noteworthy observation arises from the additional information about the synonymous verbs ‘restrict’ and ‘constrain’, discovered in the online version of LDOCE. Despite their near-synonymous nature, the provided sentence examples, which were drawn from the LDOCE’s corpus itself, demonstrate an observable difference in the frequency of active and passive constructions between ‘restrict’ and ‘constrain’. To obtain a larger sample size, the researcher consulted COCA to randomly generate 100 concordance lines where the near-synonyms ‘restrict’ and ‘constrain’ were specified as verbs. The results are presented in Table 2.

Table 2

Distributions of Passive and Active Constructions of ‘Restrict’ and ‘Constrain’ in LDOCE Online and COCA

Sources	Restrict		Constrain	
	Active	Passive	Active	Passive
LDOCE	87.50%	12.50%	33.33%	66.67%
COCA	77.00%	23.00%	44.00%	56.00%

Table 2 illustrates the usage of ‘restrict’ in the online version of LDOCE and randomly generated concordance lines in COCA. In LDOCE Online, 7 sentence examples, accounting for 87.50% of the examples, are presented in the active voice, and only 1 in the passive voice, equaling 12.50%. Similarly, in COCA, 77.00% and 23.00% of the randomized concordance lines have ‘restrict’ as active and passive verbs, respectively. In contrast, the distributions of ‘constrain’ indicate that the verb is more frequently used in the passive voice, accounting for 66.67% of the examples in LDOCE and

56% of the automated concordance lines from COCA, compared to 33.33% and 44% of active constructions discovered in the online version of LDOCE and COCA, respectively. This similar trend from both sources suggests that there may be preferred grammatical constructions for each of these near-synonyms. Specifically, ‘constrain’ tends to be used more passively and less actively than ‘restrict’ and vice versa. The examples of ‘constrain’ in passive constructions, elicited from concordance lines randomly generated by COCA are given below.

- (1) Thus, *youth are constrained* at both ends of a witchcraft incident.
- (2) *Ethanol prices could also be constrained* by infrastructure that is....
- (3) When you have such limited means, *you are constrained* to use them in a limited manner....
- (4) *NASA programs are often constrained* by limits of production, and physical limits that....
- (5) But *all were constrained* to honor god publicly or else....
- (6) Within the first two years *the overturning is strongly constrained* by the ocean’s observed density field.
- (7) During testing, *each insert was constrained* between two 19-cm by 19-cm aluminum plates, spaced 10.00 mm apart....
- (8) In other words, it is assumed that *performance will not be constrained* by limited vocabulary skills.
- (9) *The best efforts* of these admirable gatekeepers *are constrained* by the logistics under which they labor.
- (10) *Application* of external support forces *was constrained* to specific locations corresponding to the foot center of mass....

This corpus-informed passivity of ‘constrain’, discovered in the current study, as in 1-10 above, can be supported by the idea that certain verbs are used more passively than actively (Alexander, 1990). In addition, although ‘constrain’ is grammatically acceptable in both active and passive constructions, its preference for passivity might be influenced by lexical factors, similar to other verbs that are passivized most of the time in sentence constructions such as *be subjected to*, *be based on*, and *be positioned* (Biber et al., 1999, p. 479). However, this information regarding the structural preferences of both near-synonyms ‘constrain’ and ‘restrict’ is not explicitly provided in the online version of LDOCE.

In the following section, the results on the distributions of the synonymous verbs ‘restrict’ and ‘constrain’ across different textual sections in COCA will be presented and discussed. Subsequently, the noun collocates of these two near-synonyms will be investigated.

Table 3

Distribution of Frequency of the Near Synonyms 'Restrict' and 'Constrain' across All Genres in COCA

Restrict			Constrain		
Genre	Frequency	Per Million	Genre	Frequency	Per Million
academic texts	1,578	13.17	academic texts	679	5.67
webpages	1,125	9.05	webpages	221	1.78
newspapers	926	7.61	blogs	162	1.26
blogs	931	7.24	magazines	140	1.11
magazines	881	6.99	newspapers	69	0.57
spoken	566	4.49	spoken	64	0.51
fiction	91	0.77	fiction	22	0.19
TV/movie	79	0.62	TV/movie	8	0.06
subtitles			subtitles		
Total	6,177		Total	1,365	

In Table 3, the total occurrence of 'restrict' is much more frequent than that of 'constrain' by around five times (6,177 vs. 1,365 tokens). Additionally, the verb 'restrict' occurs with greater frequency compared to the verb 'constrain' in all textual sections of COCA. Regarding formality, the occurrences of 'restrict' and 'constrain' are most prevalent in the section of academic texts (13.17 and 5.67 per million, respectively), firmly establishing their exclusive association with academic genres. The results of this study, which emphasize the highly academic nature of both 'restrict' and 'constrain', align with Coxhead's AWL (2000), of which both of these near synonyms are members. Furthermore, when considering the list of textual sections where the synonyms 'restrict' and 'constrain' occur, the sequences are nearly identical. The only difference is that, that sorted in descending orders of frequency, the rankings between third and fifth are slightly different. The occurrences of 'restrict' ranked third (7.61 per million), fourth (7.24 per million), and fifth (6.99 per million) in the sections of newspapers, blogs and magazines, respectively, whereas the occurrences of 'constrain' ranked third (1.26 per million), fourth (1.11 per million), and fifth (0.57 per million) in blogs, magazines and newspapers, respectively. These nearly identical sequences, where the sections of academic texts, webpages, blogs, magazines, and newspapers reveal higher occurrences of 'restrict' and 'constrain', strengthen the indisputable role of this pair of near-synonyms in formal English, particularly in written genres. In addition, based on the corpus-informed data, the lower occurrences of 'restrict' and 'constrain' in the sections of spoken, fiction, and television and movie subtitles indicate their

detachment from informal language and, conversely, their association with formal English.

Table 4

Noun Collocates of 'Restrict' and 'Constrain' from COCA

Rank	Noun collocate	Restrict		Noun collocate	Constrain	
		Frequency	MI Score		Frequency	MI Score
1	access	372	7.93	ability	24	6.15
2	use	171	3.35	behavior	24	5.97
3	law*	117	3.72	power	21	4.39
4	ability	110	6.17	growth	18	5.9
5	abortion	102	7.48	choice	17	5.27
6	freedom	101	6.53	action	17	4.96
7	speech	75	6.02	development	15	4.88
8	movement	69	5.55	rule*	15	4.81
9	activity	58	4.84	president	14	3.22
10	number	58	3.38	freedom	12	5.63
11	regulation*	53	5.28	government	12	3.33
12	immigration	52	6.72	model	11	4.28
13	flow	46	5.87	proportion	10	7.31
14	power	46	3.34	activity	10	4.48
15	trade	35	4.47	discretion	8	8.69
16	gun	35	4.37	cost	8	3.57
17	voting	34	6.74	scope	7	6.79
18	sale	34	6.10	parameter	6	7.09
19	amount	34	4.32	decision-making	5	7.75
20	choice	34	4.09	capacity	5	5.32
21	import	31	6.87	possibility	5	4.71
22	sales	31	5.07	theory	5	4.08
23	liberty	30	6.11	box	5	3.97
24	travel	29	4.18	judge	5	3.65
25	export	28	6.57	creativity	4	6.65
26	scope	27	6.56	emission	4	5.94
27	development	27	3.55	expansion	4	5.79
28	food	27	3.05	spending	4	5.09
29	competition	26	5.17	estimate	4	4.29
30	growth	26	4.25	exercise	4	4.17

*Left-sided noun collocates

Table 4 presents the top-30 noun collocates for 'restrict' and 'constrain' with MI scores equal to or exceeding 3.00. Arranged in descending order of frequency, each of these lemmatized noun collocates encompasses occurrences of either singular or plural forms the target synonymous verbs. Both 'restrict' and 'constrain' have a certain degree of synonymy since these two verbs manually share nearly one-third of the total number of noun

collocates that made their way into the top 30 of this frequency list. The eight nouns that are mutually collocated with ‘restrict’ and ‘constrain’ consist of *ability*, *activity*, *choice*, *development*, *freedom*, *growth*, *scope*, and *power*. All of these nouns are the right collocates of ‘restrict’ and ‘constrain’, and they are associated with the two synonyms in a verb-object pattern. Moreover, among these shared noun collocates, *ability*, *activity*, *freedom*, and *power* appear to be the most interchangeable between these two near synonyms, as they are ranked in the top half of the frequency lists. This is particularly evident with *ability*, which holds both the fourth and first positions in the descending order of occurrences of noun collocates of ‘restrict’ and ‘constrain’, as illustrated below.

- (11) But the law does not specifically **restrict the ability** of Americans to collect and store seeds for their own purposes.
- (12) International legal principles, including respect for another nation's sovereignty, **constrain our ability** to act unilaterally.

The interchangeability between the synonymous verbs ‘restrict’ and ‘constrain’ is evident, as seen in 11 and 12. Furthermore, both synonyms appear in similar contexts, where law enforcement and authority prevent *ability*, as indicated by *the law* and *international legal principles*, respectively.

- (13) Current gaps in taxonomic knowledge and expertise **restrict our ability** to effectively conserve and manage biodiversity.
- (14) Changes in moisture could **restrict its ability** to survive without irrigation.
- (15) Career-advancement issues related to research and publication **constrain the ability** of junior faculty to go abroad.
- (16) The pressures on kids today also **constrain their ability** to play and use their imaginations.

In 13-14, the verb ‘restrict’ is present in contexts where *ability* is conditioned not only by laws but also by external forces, as evidenced by *current gaps* and *changes in moisture*. Similar to 15-16, outside influences, like *career-advancement issues* and *the pressures on kids*, curtail *ability* through the verb ‘constrain’.

All noun collocates of the synonymous verbs ‘restrict’ and ‘constrain’ in Table 4 will now be categorized into thematic groups with similar semantic preferences, as subsequently shown in Tables 5-6, respectively.

Table 5*Semantic Preference of Noun Collocates of 'Restrict'*

1. COMMERCE	amount, competition, export, import, number, sale, sales, trade
2. RIGHTS	abortion, choice, freedom, liberty, speech, voting
3. GOVERNANCE	gun, law, power, regulation
4. MOTION	flow, movement, travel
5. POTENTIAL	ability, scope, use
6. PROGRESS	development, growth
7. INWARDNESS	access, immigration
8. PERFORMANCE	activity
9. MISCELLANEOUS	food

In Table 5, the analysis of noun collocates reveals nine emerging themes of semantic preferences for 'restrict', which include COMMERCE, RIGHTS, GOVERNANCE, MOTION, POTENTIAL, PROGRESS, INWARDNESS, PERFORMANCE, and MISCELLANEOUS. The majority of the noun collocates of the verb 'restrict' can be categorized into COMMERCE (i.e., *amount, competition, export, import, number, sale, sales, trade*), RIGHTS (i.e., *abortion, choice, freedom, liberty, speech, voting*), GOVERNANCE (i.e., *gun, law, power, regulation*), emphasizing the three central themes associated with the verb 'restrict', as exemplified in 17-19.

- (17) The members would **restrict inter-Nordic trade** to protect their respective home markets.
- (18) These are just a few examples, but we do **restrict freedom** of speech for various reasons.
- (19) Liberals and libertarians have complained that this sentence would radically **restrict the powers** of state legislatures.

Other thematic variations of noun collocates, which are semantically preferred by the verb 'restrict' entailed MOTION (i.e., *flow, movement, travel*, as in 20), POTENTIAL (i.e., *ability, scope, use*, as in 21), PROGRESS (i.e., *development, growth*, as in 22), INWARDNESS (i.e., *access, immigration*, as in 23), and PERFORMANCE (i.e., *activity*, as in 24). The noun *food*, which does not have semantic association with any of the aforementioned themes, is categorized as MISCELLANEOUS.

- (20) They squeeze the carotid artery and **restrict blood flow** to the occipital lobe.

- (21) The presence of civilians can **restrict the use** of fires and reduce the combat power available to a platoon leader.
- (22) Many of these cloud platforms **restrict software development** to little more than piecing together existing modules to form an app.
- (23) More governments could **restrict access** to Facebook.
- (24) Many seniors experience fear of falling and **restrict their activities**.

Table 6*Semantic Preference of Noun Collocates of 'Constrain'*

1. GOVERNANCE	government, judge, power, president, rule
2. RIGHTS	choice, decision-making, discretion, freedom
3. POTENTIAL	ability, capacity, possibility, scope
4. PERFORMANCE	action, activity, behavior, exercise
5. INVENTION	creativity, model, theory
6. MEASUREMENT	estimate, parameter, proportion
7. PROGRESS	development, growth
8. OUTWARDNESS	emission, expansion
9. COMMERCE	cost, spending
10. MISCELLANEOUS	box

Table 6 presents ten unfolding themes of noun collocates for 'constrain'. The four key semantic preferences of the verb 'constrain' include GOVERNANCE (i.e., *government, judge, power, president, rule*, as in 25), RIGHTS (i.e., *choice, decision-making, discretion, freedom*, as in 26), POTENTIAL (i.e., *ability, capacity, possibility, scope*, as in 27), and PERFORMANCE (i.e., *action, activity, behavior, exercise*, as in 28).

- (25) Statutes cannot **constrain a president** who will not be constrained.
- (26) Domination can occur when powerful groups control political agendas and **constrain the choices** of the less powerful.
- (27) The latter **constrains the capacity** of peasant farmers to reorganize production.
- (28) These realities shape and **constrain human behavior**.

Other thematic variations of noun collocates semantically associated with the verb 'constrain' are *estimate, parameter, and proportion*, which are thematized as MEASUREMENT, as in 29, and *creativity, model, and theory*, which are categorized as INVENTION, as in 30. Additional groupings of semantic preference for the verb 'constrain' consist of PROGRESS (i.e., *development, growth*, as in 31), COMMERCE (i.e., *cost, spending*, as in 32), and

OUTWARDNESS (i.e., *emission*, *expansion*, as in 33). Semantically disassociated from other groupings, the noun *box* is assigned to MISCELLANEOUS.

- (29) Taking into account these observations we used a noisy Gompertz curve to **constrain the parameters** of our tumor growth model.
- (30) That incredible sensitivity underlines how such bounds can **constrain models** of new physics.
- (31) Some physicists think that Moore's Law may break down in the near future and **constrain the growth** of computing power.
- (32) Promising innovations, which **constrain costs** and increase completion, while improving quality, are still the exception today.
- (33) DECam will observe type Ia supernova and baryon acoustic oscillations and this will be to **constrain the expansion** of the universe.

Out of these ten groupings of semantic preferences, six overlap with those of 'restrict'. The semantic themes, mutually shared by the verb 'constrain' and its synonymous counterpart are as follows: GOVERNANCE, RIGHTS, POTENTIAL, PERFORMANCE, PROGRESS, and COMMERCE.

It is important to note that the lists include both left and right collocates for 'restrict' and 'constrain'. The left collocates are associated with the target synonyms in a subject-verb manner, while the right collocates are in verb-object fashion. However, as cautioned by Phoocharoensil (2021a), the list of collocates, generated natively by COCA, could be misleading due to the inclusion of false collocates. Therefore, it should be interpreted with caution, and researchers should exercise discretion when reviewing the original list. Similarly, in this study, not all nouns, categorized as collocates by COCA, are genuine noun collocates of 'restrict' and 'constrain' because they lack direct semantic relationships with the target synonyms, especially in the subject-verb pattern. In fact, many nouns, considered by COCA as left collocates, are simply in close proximity to 'restrict' and 'constrain' without functioning as nominal subjects of the synonymous verbs.

An illustrative example is *effort*, which COCA initially ranked third in the frequency list of noun collocates for 'restrict', with a total frequency of 119 and a MI score of 4.72. However, upon closer examination, it becomes evident that more than 95% of the concordance lines where *effort* is situated on the left-hand side of 'restrict' do not represent a subject-verb relationship between the two words. For instance, in the concordance line *Computer companies succeeded in weakening efforts to restrict overseas sales of software that encrypts data*, the noun *effort* does not function as the subject of the verb 'restrict', and

conversely, the verb ‘restrict’ is not a predicate of the noun *effort*. Consequently, nouns of this type, mistakenly identified by COCA as left collocates of the target synonyms, have been excluded from the analysis.

This explains why there are only 2 left noun collocates for ‘restrict’ and 1 for ‘constrain’ (*law*, *regulation*, and *rule*, respectively) since the subject-verb relationship between other left noun collocates, apart from the three aforementioned, and the target synonyms is not clearly established. Additionally, although in many concordance lines the left noun collocates—*law*, *regulation*, and *rule*—are not the subjects of the sentences, they are, indeed, the subjects of the relative clauses with verbs ‘restrict’ and ‘constrain’, as exemplified in the following concordance line:...*that he will not accept any abortion laws that restrict the current status quo*. In this case, *laws* are considered left noun collocates since they are the subjects of the verbs ‘restrict’ within the relative clause, implying what the laws enforce.

Even though the left noun collocates are less frequently associated with the two target synonyms, as compared to the right noun collocates, further examinations were intentionally performed by the researcher. This is because—collectively—*law*, *regulation*, and *rule* belong to the same semantic group of GOVERNANCE. Examining the concordance lines in COCA where *law*, *regulation*, and *rule* precede the target synonyms ‘restrict’ and ‘constrain’ in a subject-verb pattern, the researcher discovered that this subject-verb structure is, in fact, followed by other right noun collocates, which are ranked in the top 30 of the frequency list. These right-sided noun collocates of ‘restrict’ and ‘constrain’, which are introduced by *law*, *regulation*, and *rule*, include *abortion*, *ability*, *power*, *use*, *access*, *trade*, and *behavior* are presented in 34-43.

- (34) According to this nice summary from Reproductive Health Reality Check ***federal and state laws severely restrict late term abortions***.
- (35) Theoretically ***RTW laws may restrict an employer's ability*** to enter a closed shop agreement.
- (36) ***Such general laws shall restrict the powers*** of such cities, towns and villages to borrow money and contract debts.
- (37) ***His law would restrict children's access*** to shows with “violent” special effects and stunts.
- (38) ***Regulations can restrict the use*** of engine braking.
- (39) Although ***the regulations restrict access*** to PHI for specific purposes, the reality is that when more people have....
- (40) ***Limited-entry regulations restrict access*** while allowing as much freedom as possible to choose how, when, and....

-
- (41) In other words, even if ***legal rules constrain the subsequent use*** of already-collected information, it can still be reasonable to worry that...
- (42) However, ***EPA trading rules inefficiently constrain intra-firm and inter-firm trade***.
- (43) ***Some moral rules constrain the behavior*** of individuals in economically efficient ways, benefiting society at large.

As can be seen in 34-43, the left-sided noun collocates, *law*, *regulation*, and *rule*, belonging to the same semantic group of GOVERNANCE, act upon those on the right of the synonymous verbs ‘restrict’ and ‘constrain’, which were semantically grouped as RIGHTS (i.e., *abortion*), POTENTIAL (i.e., *ability* and *use*), INWARDNESS (i.e., *access*), COMMERCE (i.e., *trade*), and PERFORMANCE (i.e., *behavior*). As these concordance lines were extracted from COCA, which is based in the U.S., the coexistences between the target synonymous verbs ‘restrict’ and ‘constrain’ and these specific left- and right-sided noun collocates, could be significantly influenced by social factors (Reiter & Sripada, 2004), which, in this specific case, are public issues central to American society.

For instance, women’s rights to receive abortions, as a result of unexpected pregnancies, has been one of the divisive and hotly debated topics in the U.S. According to statistical data collected by the Centers for Disease Control and Prevention (CDC), approximately 620,000 American women underwent abortions in the year 2020. Even though it was estimated that more than one-fourth of the American female population would have abortions before they turned 45 years of age, there had been a number of legal attempts by the U.S. government to limit women’s access to abortions in the past (Jones & Jerman, 2014). This has been seen by many as a violation of human rights, committed by law enforcers.

Not only does this instance create a highly context-specific environment where the two synonymous verbs ‘restrict’ and ‘constrain’ occur, but it also constructs a complete sentential pattern of subject-verb-object (S-V-O). This is graphically summarized in Figure 1 below.

Figure 1

Diagrammatic Representation of the Linkage between ‘Restrict’ and ‘Constrain’ and Left- and Right-sided Noun Collocates in S-V-O Pattern

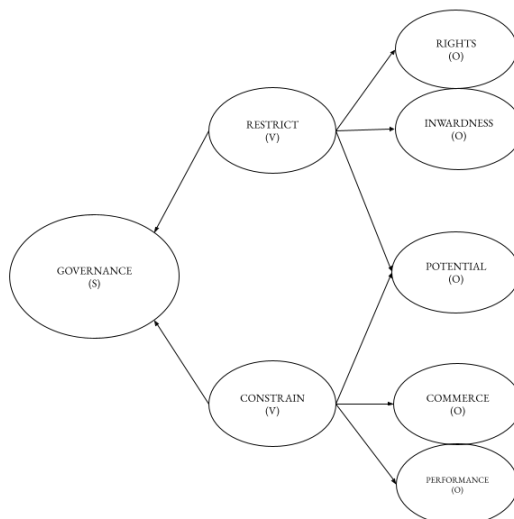


Figure 1 illustrates the corpus-informed sequences, where the target synonymous verbs ‘restrict’ and ‘constrain’ act as intermediaries between the left noun collocates, semantically representing GOVERNANCE (i.e., *law*, *regulation*, and *rule*) and right noun collocates, grouped by semantic preferences of RIGHTS, POTENTIAL, INWARDNESS, COMMERCE, and PERFORMANCE.

Despite their similarities in superficial meaning, subtle semantic differences are observable. In this specific context, the subject collocates, *law* and *regulation*, have ‘restrict’ as verbs, acting upon its object collocates, which semantically imply RIGHTS, POTENTIAL, and INWARDNESS, while *rule* has ‘constrain’ as the verb, which is followed by its object collocates, denoting COMMERCE, POTENTIAL, and PERFORMANCE.

As can be seen, the only semantic group of right-sided noun collocates, POTENTIAL, is shared by the two synonymous verbs. This implies that, even in this very specific context where ‘restrict’ and ‘constrain’ share the left-sided noun collocates of a similar semantic preference (i.e., GOVERNANCE), the two synonymous verbs, however, may require other right noun collocates of different semantic preferences: which consist of RIGHTS and INWARDNESS for ‘restrict’ and COMMERCE and PERFORMANCE for ‘constrain’.

This is supported by the presence of other right-sided noun collocates with MI scores higher than 3.00 in the top 30 list. The listed right-sided noun collocates of the two verbs, such as *liberty* (MI=6.11), and *immigration* (MI=6.72), appear as direct objects in the S-V-O pattern of GOVERNANCE-RESTRICT-RIGHTS/INWARDNESS, as in 44 and 45, meanwhile *spending* (MI=5.09) and *action* (MI=4.96), function as direct objects in the S-V-O construction, representing GOVERNANCE-CONSTRAIN-COMMERCE/PERFORMANCE, as in 46-47.

- (44) ***Laws and regulations that restrict individual liberties*** are routinely enacted to protect and promote public health and welfare.
- (45) However, ***this legislation did not restrict immigration*** from Mexico and other countries in the Western Hemisphere.
- (46) Fewer states have ***rules that constrain capital spending*** and trust funds (33 and 30 states, respectively).
- (47) ***Tobacco-specific laws constrain only some actions*** of the tobacco industry.

Since concordance lines 44-47 are data obtained from COCA, which is based in the U.S., they again convey meanings related to the country's diverse societal issues. This suggests that, in addition to semantic preferences, contextual influences (e.g. social factors, collocation) may also play a vital role in the choice between synonymous words (Reiter & Sripada, 2004).

Another revelation worth discussing is the stark contrast in morphological meanings between the two semantic groupings of the right-sided collocates of the synonyms 'restrict' and 'constrain': INWARDNESS and OUTWARDNESS. As previously demonstrated, the object noun collocates of the verb 'restrict', semantically belonging to INWARDNESS, are *access* and *immigration*, whereas those of the verb 'constrain', categorically assigned to OUTWARDNESS, include *emission* and *expansion*. Even though noun collocates in the former grouping are not antonyms of those in the latter, it is relatively obvious that each noun collocate in the semantic groups of INWARDNESS and OUTWARDNESS is composed of morphological units representing inbound (i.e., *acc-* in *access* and *im-* in *immigration*) and outbound (i.e., *e-* in *emission* and *ex-* in *expansion*) activities, respectively.

This is further evidenced by other nouns with MI scores higher than 3.0 that, however, fall outside of the top 30 frequency list of noun collocates of 'restrict' and 'constrain'. In addition to those in the lists, 'restrict' also collocates with *intake* (MI=7.11) and *entry* (MI=5.70), conceivably grouped as INWARDNESS, as in 48 and 49, while 'constrain' is associated with *outlay* (MI=7.96), and *emergence* (MI=7.11), potentially categorized as OUTWARDNESS, as in 50-51.

- (48) Rather than restrict weekend calories, one should **restrict food intake** during the week and then allow certain “treats” on the weekend.
- (49) In one, bacteria **restrict the entry** of antibiotics into the cell; in the other, bacteria produce an....
- (50) The first thing you need to do is **constrain public outlays** for the old.
- (51)authoritarian regimes at some point will **constrain the emergence** of a market economy.

As can be seen in 48-49, the right-sided noun collocates of ‘restrict’, excluded from the top-30 frequency list, share morphemes representing internality with those included in INWARDNESS (i.e., *in-* in *intake* and *en-* in *entry*), while those of ‘constrain’, which were also discarded, share lexical constituents indicative of externality with the object collocates of ‘constrain’ (i.e., *out-* in *outlay*, and *e-* in *emergence*), as in 50-51. This difference between the morphological meanings of the object collocates of the two synonymous verbs ‘restrict’ and ‘constrain’ is illustrated in Figure 2 below.

Figure 2

Diagrammatic Representation of the Opposite Directions of Morphological Meanings in Right-sided Noun Collocates of ‘Restrict’ and ‘Constrain’



In Figure 2, ‘restrict’ is positioned at one end of the spectrum, while ‘constrain’ is at the other end. The verb ‘restrict’ tends to convey a sense of limiting something from moving toward the inside since it is associated with object collocates that are partly formed by morphemes indicating an inward direction (i.e., *acc-* in *access*, *im-* in *immigration*, *in-* in *intake*, and *en-* in *entry*). In contrast, the verb ‘constrain’ is more likely to mean hindering something from going out, as it is collocated with nouns that are comprised of lexical units representing an outward direction (i.e., *e-* in *emission*, *ex-* in *expansion*, *out-* in *outlay*, and *e-* in *emergence*).

Conclusion and Pedagogical Implications

Emerging from the COCA-informed data, both synonymous verbs ‘restrict’ and ‘constrain’ appear to have the highest occurrences in academic genres, solidifying their presence in formal English and Coxhead’s (2000) AWL. The synonymous pair also shares six overlapping groups of noun collocates, characterized by semantic preference, consisting of GOVERNANCE, RIGHTS, POTENTIAL, PERFORMANCE, PROGRESS, and COMMERCE. However, ‘restrict’ is associated with an additional semantic group of noun collocates (i.e., INWARDNESS), whereas ‘constrain’ features three other distinctive groups (i.e., MEASUREMENT, INVENTION, and OUTWARDNESS).

Interestingly, despite having far more tokens in the corpus, ‘restrict’ exhibits a narrower semantic range of noun collocates, in comparison with ‘constrain’. The observed pattern of both near-synonyms co-existing with two groups of right-sided noun collocates that contain completely opposite morphological meanings, i.e., *acc-* in *access*, as in *restrict access*, versus *ex-* in *expansion*, as in *constrain expansion*, could significantly impact how ELT practitioners approach these two synonymous verbs in their instruction. Based on this, teachers should make students aware of the concept of near-synonyms and the fact that complete substitutability between them across contexts is almost impossible (Jackson & Amvela, 2007; Phoocharoensil, 2022). Furthermore, both synonymous verbs appear to be highly context-dependent and subject-specific. According to the corpus data, they are exclusively associated with a specific group of left-sided noun collocates, semantically representing GOVERNANCE in subject-verb patterns (i.e., *laws* or *regulations restrict* and *rules constrain*). As ‘restrict’ and ‘constrain’ are part of Coxhead’s (2000) AWL, they are considered specialized vocabularies (Nation, 2001) that are contextually associated with law and commerce. Teaching these two synonymous verbs, along with the findings of this corpus-based study, might help learners of English, who study law and commerce, become more aware of the subtle differences in authentic usage between them, in terms of collocations and semantic preferences. This is because instructional activities involving the usages of technical vocabularies will be particularly meaningful for students from specific academic disciplines (Nation, 2001).

However, according to Nation (2001, p. 20), “one person’s technical vocabulary is another person’s low-frequency word”. This, consequently, may raise a thought-provoking question for ELT practitioners regarding the necessity of including the synonymous and technical AWL verbs ‘restrict’ and ‘constrain’ in the instruction since generic learners of the English language, who are not enrolled in academic courses in English for Specific Purposes, may perceive them as surplus to requirements.

The results of this corpus-based study seem to have a few limitations. First, this study was limited to extracting the top-30 noun collocates of the target synonyms using MI scores and categorizing them based on semantic preferences. It is expected that a higher number of noun collocates will lead to the emergence of new themes associated with the two synonymous verbs. Secondly, as contextual variables could influence the coexistence between the target synonyms and their collocations (Reiter & Sripada, 2004), the findings of this study could be influenced by underlying social factors because they are derived from the consultation with COCA, which is American-based. Using other established corpora (e.g., The British National Corpus or BNC) might produce different results. Lastly, this study utilized two primary criteria for distinguishing synonyms: level of formality and collocations. The inclusion of other synonym-discriminating criteria (e.g., colligations) would provide more comprehensive insights regarding the target synonyms.

Acknowledgements

I would like to express my sincere gratitude to Assoc. Prof. Dr. Supakorn Phoocharoensil, Editor-In-Chief of this prestigious journal, for his invaluable guidance and constructive feedback in bringing this post-doctoral research of mine to fruition.

An honorable mention also goes to Ajarn Oliver Oonlamai of the Faculty of Humanities of Phetchabun Rajabhat University, Thailand, for his English native speaker's intuition in the process of semantic groupings of the noun collocates of the target synonyms and proofreading this paper.

About the Author

Nateethorn Narkprom: A lecturer of English in the Faculty of Education of Phetchabun Rajabhat University, Thailand. His research interests involve corpus linguistics and textual analysis.

References

- Ajmal, M., Kumar, T., Ritonga, M., & Nukapangu, V. (2022). A corpus-based analysis of the adjectives and synonyms-beautiful, handsome, and pretty. *World Journal of English Language*, 12(2), 159–168. <https://doi.org/10.5430/wjel.v12n2p159>
- Alexander, L. G. (1990). *Longman English grammar practice*. Pearson Education Limited.
- Aroonmanakun, V. (2015). Quick or fast: A corpus-based study of English synonyms. *LEARN Journal: Language Education and Acquisition Research Network*, 8(1), 53–62. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/102685>
- Aroonmanakun, V., & Aroonmanakun, W. (2023). A comparison of 'little' and 'small' in English and Thai: A corpus-based study. *LEARN Journal: Language Education and Acquisition Research Network*, 16(1), 296–319. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/263443>
- Baker, P., Hardie, A., & McEnery, T. (2006). *A glossary of corpus linguistics*. Edinburgh University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education Limited.
- Bennett, G. R. (2010). *Using corpora in the language learning classroom*. University of Michigan Press. <https://doi.org/10.3998/mpub.371534>
- Boontam, P., & Phoocharoensil, S. (2022). Broaden your horizons: Distribution and collocational patterns of the English synonyms "expand," "widen," and "broaden." *The International Journal of Communication and Linguistic Studies*, 20(1), 107–123. <https://doi.org/10.18848/2327-7882/CGP/v20i01/107-123>
- Carter, R. (2012). *Vocabulary: Applied linguistic perspectives*. Routledge.
- Chaengchenkit, R. (2023). A corpus-based study of the synonyms "cease," "halt," and "stop." *LEARN Journal: Language Education and Acquisition Research Network*, 16(1), 473–494. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/263452>
- Chaokongjakra, W. (2023). Is it 'important,' 'significant,' or 'crucial'? A corpus-based study of English synonyms. *LEARN Journal: Language Education and Acquisition Research Network*, 16(2), 512–532. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/266969>
- Cheng, W. (2012). *Exploring corpus linguistics: Language in action*. Routledge.
- Choo, L. B., Lin, D. T. A., Singh, M. K. M., & Ganapathy, M. (2017). The significance of the Academic Word List among ESL tertiary

- students in a Malaysian public university. *3L: The Southeast Asian Journal of English Language Studies*, 23(4), 56–65.
<http://dx.doi.org/10.17576/3L-2017-2304-05>
- Clark, E. V. (1992). Conventionality and contrast: Pragmatic principles with lexical consequences. In A. Lehrer & E. F. Kittay (Eds.), *Frames, fields, and contrasts: New essays in semantic and lexical organization* (pp. 171–188). Lawrence Erlbaum Associates, Inc.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238.
- Cruse, A. (1986). *Lexical semantics*. Cambridge University Press.
- Crystal, D. (2003). *A dictionary of linguistics and phonetics*. Blackwell Publishing.
- Edmonds, P., & Hirst, G. (2002). Near-synonymy and lexical choice. *Computational Linguistics*, 28(2), 105–144.
<https://doi.org/10.1162/089120102760173625>
- Flowerdew, L. (2012). *Corpora and language education*. Palgrave Macmillan.
- Folse, K. (2011). Applying L2 lexical research findings in ESL teaching. *TESOL Quarterly*, 45(2), 362–369.
- Imsa-ard, P. (2021). A corpus-based study on the meanings, distribution, collocations, and formality of 'concentrate,' 'emphasize,' and 'focus'. *NIDA Journal of Language and Communication*, 26, 22–40.
- Imsa-ard, P., & Phoocharoensil, S. (2022). "A whole new world... Wait, is it a whole, entire, or total world?": The extraction of collocations for the three English synonym discrimination. *3L: Language, Linguistics, Literature*, 28(2). <http://doi.org/10.17576/3L-2022-2802-05>
- Jackson, H., & Amvela, E. (2007). *Words, meaning, and vocabulary: An introduction to modern English lexicology*. Cassel.
- Jarunwaraphan, B., & Mallikamas, P. (2020). A Corpus-based study of English synonyms: Chance and opportunity. *REFlections*, 27(2), 218–245. <https://so05.tci-thaijo.org/index.php/reflections/article/view/248710>
- 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. <https://so06.tci-thaijo.org/index.php/thoughts/article/view/163389>
- Jones, R. K., & Jerman, J. (2014). Abortion incidence and service availability in the United States, 2011. *Perspectives on sexual and reproductive health*, 46(1), 3–14. <https://doi.org/10.1363/46e0414>
- Kruawong, T., & Phoocharoensil, S. (2022). A genre and collocational analysis of the near-synonyms teach, educate and instruct: A corpus-based approach. *TEFLIN Journal*, 33(1), 75–97.
<https://doi.org/10.15639/teflinjournal.v33i1/75-97>

- Laufer, B. (1990). Words you know: How they affect the words you learn. In J. Fisiak (Ed.), *Further insights into contrastive linguistics* (pp. 573–593). John Benjamins.
- Lei, L., & Liu, D. (2018). The academic English collocation list: A corpus-driven study. *International Journal of Corpus Linguistics*, 23(2), 216–243. <https://doi.org/10.1075/ijcl.16135.lei>
- Lertcharoenwanich, P., & Phoocharoensil, S. (2022). A corpus-based study of the near-synonyms: purpose, goal and objective. *REFLECTIONS*, 29(1), 148–168. <https://so05.tci-thaijo.org/index.php/reflections/article/view/258872>
- Lindquist, H. (2009). *Corpus linguistics and the description of English*. Edinburgh University Press.
- Liu, D. (2010). Is it a chief, main, major, primary, or principal concern?: A corpus-based behavioral profile study of the near-synonyms. *International Journal of Corpus Linguistics*, 15, 56–87. <https://doi.org/10.1075/ijcl.15.1.03liu>
- Martin, M. (1984). Advanced vocabulary teaching: The problem of synonyms. *The Modern Language Journal*, 68(2), 130–137. <https://doi.org/10.2307/327139>
- Moon, R. (2010). What can a corpus tell us about lexis?. In A. O’Keefe & M. McCarthy, M. (Eds.), *The Routledge handbook of corpus linguistics* (pp.197–211). Routledge.
- Murphy, M. L. (2009). *Semantic relations and the lexicon*. Cambridge University Press.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Petcharat, N., & Phoocharoensil, S. (2017). A corpus-based study of English synonyms: Appropriate, proper, and suitable. *LEARN Journal: Language Education and Acquisition Research Network*, 10(2), 10–24. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/111700>
- Phoocharoensil, S. (2020a). A genre and collocational analysis of consequence, result, and outcome. *3L: Southeast Asian Journal of English Language Studies*, 26(3). <https://ejournals.ukm.my/3l/article/view/38851>
- Phoocharoensil, S. (2020b). Collocational patterns of the near-synonyms error, fault, and mistake. *The International Journal of Communication and Linguistic Studies*, 19(1), 1. <https://doi.org/10.18848/2327-7882/CGP/v19i01/1-17>
- Phoocharoensil, S. (2021a). Semantic prosody and collocation: A corpus study of the near-synonyms persist and persevere. *Eurasian Journal of Applied Linguistics*, 7(1), 240–258. <https://doi.org/10.32601/ejal.911269>

- Phoocharoensil, S. (2021b). Multiword units and synonymy: Interface between collocations, colligations, and semantic prosody. *GEMA Online Journal of Language Studies*, 21(2), 28–45.
<https://ejournal.ukm.my/gema/article/view/45200>
- Phoocharoensil, S. (2022). Primary, main, and major: Learning the synonyms through corpus data. *GEMA Online Journal of Language Studies*, 22(4), 76–89. <http://doi.org/10.17576/gema-2022-2204-05>
- Phoocharoensil, S. & Kanokpermpoon, M. (2021). Distinguishing the near-synonyms *increase* and *rise*: Genre and collocation investigation. *Kasetsart Journal of Social Sciences*, 42, 968-975
- Reiter, E., & Sripada, S. (2004). Contextual influences on near-synonym choice. In A. Belz, R. Evans, & P. Piwek (Eds.), *Natural language generation* (pp. 161–170). Springer. https://doi.org/10.1007/978-3-540-27823-8_17
- Saito, K. (2020). Multi- or single-word units? The role of collocation use in comprehensible and contextually appropriate second language speech. *Language Learning*, 70(2), 548–588.
<https://doi.org/10.1111/lang.12387>
- Sangeekaew, K. (2023). A corpus-based study of English synonyms: normal, regular, common, ordinary, and usual. *The New English Teacher*, 17(1), 113–135.
- Song, Q. (2021). Effectiveness of corpus in distinguishing two near-synonymous verbs: ‘damage’ and ‘destroy’. *English Language Teaching*, 14(7), 8–20. <https://doi.org/10.5539/elt.v14n7p8>
- Sridhanyarat, K., & Phoocharoensil, S. (2023). A corpus-based investigation of English near-synonyms: Assess, evaluate, and measure. *Humanities, Arts and Social Sciences Studies*, 23(1), 208–219.
<https://doi.org/10.14456/hasss.2023.19>
- Sumonsriworakun, P. (2022). A corpus-based investigation of English synonyms: Disadvantage, downside, and drawback. *LEARN Journal: Language Education and Acquisition Research Network*, 15(2), 649–678.
<https://so04.tci-thaijo.org/index.php/LEARN/article/view/259944>
- Supanfai, P. (2022). People or persons?: A corpus-based study. *REFLECTIONS*, 29(3), 603–620. <https://so05.tci-thaijo.org/index.php/reflections/article/view/262288>
- Yang, B. (2016). A corpus-based comparative study of learn and acquire. *English Language Teaching*, 9(1), 209–220.
<https://doi.org/10.5539/elt.v9n1p209>
- Zwiers, J. (2008). *Building academic language*. International Reading Association.