A Corpus-Based Genre and Collocational Study of the Near-Synonyms: *Grasp*, *Capture*, *Seize*, *Snatch*, and *Take*

Sani Yantandu Uba[1] & Faiza Al-Dhahli[2]

[1] Department of English Language and Literature, Dhofar University, Salalah, Oman

[2] Department of English, University of Technology and Applied Sciences, Nizwa, Oman

Correspondence: Sani Yantandu Uba, Department of English Language and Literature, Dhofar University, Salalah, Oman. E-mail: suba@du.edu.om

Received: January 25, 2025 Accepted: May 16, 2025 Online Published: July 18, 2025

doi:10.5430/wjel.v16n1p73 URL: https://doi.org/10.5430/wjel.v16n1p73

Abstract

This study investigated the genre, collocational, and semantic preferences of the near-synonymous English verbs *grasp, capture, seize, snatch,* and *take.* The researchers drew data from the Corpus of Contemporary American English (COCA) and three traditional non-corpus-based dictionaries. The results revealed that writers use the verb *take* much more commonly across all eight genres, with a frequency of 863,996 out of 909,634 total occurrences of all verbs across these genres. Non-formal genres, such as TV/movies and spoken language, show the highest frequencies of the verb *take*. Additionally, the findings indicate that categorizing their adverb and noun collocates according to semantic preferences provides insights into their co-occurring contextual usage. For instance, the noun collocates of the verb *take* fall into semantic themes that include *responsibility, action, observation/perception, benefits, life, location,* and *evaluation.* In contrast, the verb ***c****apture,* which has the second highest frequency, is associated with semantic themes such as *visual representation, abstract concepts, data, focus, force, and body parts*. The third most frequent verb, *grasp,* has semantic themes related to *body parts, abstract concepts, physical objects, mental attributes, and opportunity*. These findings could help ESL/EFL teachers design lessons that focus on genre-specific language use. For instance, ESL learners could develop their ability to identify verb preferences across various genres. Additionally, COCA is freely accessible online. Teachers could engage students in task-based learning. For example, they can ask students to generate a collocational distribution list of a set of near-synonyms and then ask group the semantic preferences of the target synonyms. The students could find subtle differences between the synonyms. Ultimately, this awareness could also improve their linguistic competence.

Keywords: COCA, collocation, genre, synonyms, semantic preferences

1. Introduction

This study investigates the genre, collocational, and semantic preferences of the near-synonymous English verbs *grasp, capture, seize, snatch, and take*. Learning vocabulary is one of the key aspects of acquiring a second language (Phoocharoensil, 2020). Sridhanyarat (2018) claims that learning synonyms proves to be challenging for second language learners. Jackson and Amvela (2007) also believe that synonyms typically present difficulties for English second language learners. Many English second language learners believe synonyms are interchangeable in every context, without affecting their structure and meaning, including pragmatic meaning. However, there are many instances where synonyms are not interchangeable in the same context and structure; this attempt to substitute such synonyms might lead to unnatural and ungrammatical production of the target language (Thornbury, 2002). Near-synonymous English verbs are quite problematic for English second language learners because they have similar connotational meanings in many contexts, but they cannot be collocationally interchangeable (Chan, 2010; Nguyen & Webb, 2017; Uba & Irudayasamy, 2023). This difficulty mainly arises when English second language learners interchange synonyms without considering their collocational patterns.

Thesauruses and dictionaries typically define and describe near-synonyms in ways that imply they are interchangeable. This presentation of synonyms in these learning resources is highly under-represented because there is a lack of comprehensive descriptions of their genre, collocational, and semantic preferences. For example, the target English synonymous verbs of this study are defined in Table 1 below by the three different dictionaries.

Table 1. Definitions and examples of *grasp, capture, seize, snatch and take* from dictionaries

Word	Longman Dictionary of Contemporary English	Merriam-Webster Dictionary	Advanced American Dictionary Oxford
Grasp	"to take and hold something firmly" e.g. "Alan grasped the handle and pulled it."	"to take or seize eagerly" **e.g. "grasp** the handle and pull"	"to take a firm hold of someone or something" *e.g. "He grasped my hand and shook it warmly."*
Capture	"to get control of a place or object that previously belonged to an enemy, during a war" e.g. "The town was captured after a siege lasting ten days."	"an act of catching, winning, or gaining control by force, stratagem, or guile " e.g. "the *capture* of the city by enemy forces"	"to take control of a place, building, etc. using force" *e.g. "The city was captured in 1941."*
Seize	"to take hold of something suddenly and violently" e.g. "Suddenly, he seized my hand."	"to possess or take by force" e.g. "The bank seized the property"	**"**to take control of a place or situation, often suddenly and violently" e.g. "*He **seized power** in a military coup"*
Snatch	'to take something away from someone with a quick, often violent, movement' e.g. 'The thief snatched her purse and ran'	"to take or grasp abruptly or hastily" **e.g. "snatched** a doughnut and ran"	"to take someone or something away from a person or place, especially by force" e.g. "*The baby was snatched from its parents' car."*
Take	'to get possession or control of something' e.g. 'enemy forces have taken the airport'	"to get into one's hands or into one's possession, power, or control" e.g. "**took** them as prisoners"	"to capture a place or person; to get control of something" *e.g. "The government has **taken control of** the company."*

It is evident that these five English verbs possess similar meanings, as defined by the dictionaries in Table 1 above. However, these dictionaries do not comprehensively describe their genre usage, collocational patterns, and semantic preferences. This lack of information could limit effective learning of the similarities and differences in their semantic preferences, genre usage, and collocational behaviors. Additionally, the rationale for selecting near-synonym verbs is the emphasis by many scholars that synonyms constitute a major source of errors in learning a second language (Alanazi, 2022; Nesselhauf, 2003; & Partington, 1998). Therefore, one strategy that could facilitate learning of the similarities and differences in their genre, collocational, and semantic preferences is a corpus-based synonym study, which proves to be effective in distinguishing near-synonyms (Alanazi, 2022; Panrat & Yanasugondha, 2024; Phoocharoensil, 2020; Putklang, et al., 2024; Uba & Irudayasamy, 2023). Hence, this study aims to investigate the genre usage, collocational patterns, and semantic preferences of English synonymous verbs, *grasp, capture, seize, snatch,* and *take,* by using a corpus-based approach.

2. Literature Review

2.1 The Meaning of Lexical Items

Many scholars have argued that the traditional approach to defining words' meaning is insufficient. For example, Firth (1957) states that the context of words usually determines their meaning. While Sinclair (1998) claims that defining the meaning of a word, only on its core meaning (paradigmatic level) may lead to erroneous observation. He suggested including a syntagmatic level view of meaning (top-down presentation). This combination of paradigmatic and syntagmatic levels can provide a comprehensive description of the meaning of words (Alanazi, 2022). Furthermore, Sinclair (1998) states that the meaning of lexical items can be classified into two components: obligatory, comprising the semantic prosody of a word and its core meaning, and three optional yet genuine categories, including semantic preferences, collocation, and colligation. Alanazi (2022) notes that Sinclair (1998, p. 14) refers to this as 'coordinated secondary choices within the item, fine-tuning the meaning and giving semantic cohesion to the text as a whole'. Additionally, Hoey (2005, p. 116), in his theory of lexical priming, states that 'I would hypothesize that all words are primed for one or more collocations, semantic associations, and colligations, even if these are on the face of it unremarkable'. There seems to be a lack of a universal or unified definition of collocations. We will discuss this in the subsequent section.

2.2 Synonyms and Near-Synonyms

A synonym is a complex linguistic term. For example, Partington (1998) defines synonyms as words that may share similar meanings but show different semantic prosody behaviors and collocational patterns. Liu (2010) posits that synonyms are complex linguistic terms with similar linguistic concepts from different contexts and perspectives. Divjak (2006, p. 21) states that synonyms are one situation described in various ways and viewed from different perspectives. This demonstrates the complexity of the concept and shows that absolute synonyms are very rare in languages; 'natural languages abhor absolute synonyms just as nature abhors a vacuum' (Cruse, 1986, p. 270). As a result, near-synonyms are words that share one or more core meanings and are commonly used (Alanazi, 2022; Uba & Irudayasamy, 2023). As noted above, near-synonyms are confusing, particularly for English second language learners, because they are not interchangeable in many contexts. Substituting them in similar contexts may lead to misinterpretation or vagueness of meanings. To understand the similarities and differences of near-synonyms, researchers must conduct an in-depth analysis of their use in various contexts. This includes the paradigmatic level (their core semantic meanings) and the syntagmatic level (such as collocations, semantic

preferences, and colligation). We will discuss collocations and semantic preferences as they relate to this study below.

2.3 Collocations

Collocation refers to the common occurrence of words appearing together in texts or discourse. It was initially introduced and defined by Firth (1957, p. 11), stating that "you shall know a word by the company it keeps". Collocation is categorized into fixed or strict combinations, such as 'see you soon' and 'how are you?', and less fixed phrases, such as 'completely different/new/free' (Alanazi, 2022). Additionally, collocations differ in the number of words in sequence, including the range of collocates, as some words may have broader collocates than others (Alanazi, 2022; & Nation, 2013). Flowerdew (2012) argues that the meaning of words does not solely depend on what they possess but also includes how they collocate with other words. One possible approach to examining the similarities and differences of synonyms is to look at their possible collocates (Phoocharoensil, 2020; Uba & Irudayasamy, 2023). Synonyms can, in some instances, be different (not interchangeable) when they typically co-occur with other words. For example, '*increase' and 'rise'* share similar core meanings. However, '*increase'* has abstract noun collocates related to *finance* and *the natural environment*. On the other hand, '*rise'* has abstract noun collocates pertaining to *finance* and *security, as well as concrete* noun collocates related to the natural environment (Uba & Irudayasamy, 2023).

One approach that determines how words collocate is the statistical measure for calculating the degree of occurrence of collocates. Sinclair (1991), cited by Alanazi (2022), states that words can be considered collocates if two or more words co-occur within a particular span of each other. Three commonly used statistical measures that scholars adopt in corpus-based studies are LogDice, MI scores, and T-score. LogDice measures the strength of collocates and is a standardized measure with a maximum value of 14. Alanazi, (2022) and Gablasova et al. (2017) believe that LogDice efficiently compares different large corpora. They emphasize that other traditional statistical measures could skew some scores when used with large corpora. T-score is another statistical measure that assesses the 'certainty of collocations' (Hunstons, 2002, p. 73). T-score provides raw frequency scores, and 'collocates with high T-scores are frequently found in language' (Alanazi, 2022, p. 5). However, many scholars argue that the T-score is biased towards corpus size, making it inappropriate to compare different corpora of varying sizes (Alanazi, 2022; & Gablasova et al., 2017). The third statistical measure is the mutual information index (MI), which is typically used to calculate the ratio between the frequency of random co-occurrence and the frequency of collocates of the two words in combination (Alanazi, 2022; & Church & Hanks, 1990). The MI score shows how strongly words correlate, and how items are associated (Gardner & Davies, 2014).

2.4 Semantic Preference

Semantic preference is another approach that distinguishes the differences and similarities of synonymous words. Sinclair (2004) defines semantic preference as lexical items that share similar semantic features and typically have restricted co-occurrences. Some scholars further state that semantic preference is characterized by lexical items commonly used, restricted, and associated with a specific identifiable area or field (Ang et al., 2017; Cheng, 2012; & Phoocharoensil, 2020). In other words, semantic preference refers to words that usually co-occur and are associated with specific identifiable semantic fields; for example, the word, *teach* has a semantic relation with education and training. This assists in understanding whether synonymous words have similar semantic relationships. It also provides more information on whether the synonymous words are interchangeable or not.

2.5 Genre Theory

Genre is "a distinctive category of discourse of any type, spoken or written, with or without literary aspiration" (Swales, 1990, p. 33). Swale's genre theory emphasizes that genre consists of socially recognized communicative events. These events shared a common purpose and conventions typically recognized and understood by members of a particular discourse community (Swales, 1990). His theory highlights how genres are integral to the discourse communities, focusing not only on structural or linguistic features of texts, but also on their communicative purpose, shared among members of a specific discourse community. He further explains that 'discourse communities' are groups of people with similar goals, a specific lexis, and intercommunication practices. Swale's genre theory is tied to three key interrelated terms: genre, discourse community, and task. The communicative purpose binds all these terms:

> *It is communicative purpose that drives the language activities of the discourse community; it is communicative purpose that is prototypical criterion for genre identity, and it is communicative purpose that operates as the primary determinant of task (Swale's, 1990, p. 10)*

In other words, genres are communicative events in which members of specific discourse communities adhere to particular conventions that are shared and agreed to, aiming to facilitate the practices and needs of the speech communities. For example, in the genre of research article, the rhetorical structure of the abstract typically may include the following: background, literature review, method, results, and conclusion/implication. Thus, academic community members must adhere to these conventions when communicating their research reports.

Building on Swale's genre theory, Bhatia (1993) claims that genre is an interplay between linguistic, professional, as well as disciplinary practices. Bhatia also emphasizes that genre goes beyond only textual organization of the text but also includes the socio-cultural contexts within which the discourse community operates. One aspect of genre development that Bhatia (2004) discusses relevant to the current study is textualization of lexico-grammar. This involves context-sensitive language use because genre theory focuses on language use guided by the conventions and communicative purposes of a specific genre. Thus, a genre and corpus-based study of near synonyms

could reveal how near-synonyms are more preferred or commonly used in a specific genre because of their relatedness with the genre's register, communicative purpose and tone. Furthermore, this lexico-grammar is also concerned with rhetorical function and lexical choice. For example, different genres may impose restrictions and expectations that could influence word choice. The genre and corpus-based near-synonyms study could uncover subtle differences in formality, connotation, semantic preference, and collocation by showing their suitability for certain genres. For example, Swales (1974) found that phrases, *a give, a certain,* and *a particula*r perform a specific rhetorical function, which is unique to a particular genre. Bhatia (1992) also discovered that nominals perform different rhetorical functions across three genres investigated: advertisements, legislative provisions, and scientific research reports. The current study is underpinned by combining this genre theory and a corpus-based study of near-synonyms, aiming to uncover how the target English near-synonymous verbs are commonly used across genres, including semantic preferences and collocational behavioral profiles.

2.6 Previous Corpus-Based Studies on Near-Synonyms

Hunston (2002) and Liu (2010) believe that a corpus-based language study is more effective in providing detailed information than traditional non-corpus-based reference materials, such as dictionaries. Several corpus-based conducted studies supported this claim. For example, Gu (2017) conducted a corpus-based study of the near-synonyms *gain* and *obtain.* The results indicated that *gain* commonly co-occurs with abstract nouns associated with positive prosody. In contrast, *obtain* common co-occurs with nouns in the passive voice. Moreover, Jirananthiporn (2018) conducted a corpus-based study on synonymous nouns: *problem* and *trouble* in COCA, examining their similarities and differences in terms of adjective and verb collocates, and the distributional pattern of frequencies across genres. The findings showed that *problem* has the highest frequency across all the genres. In contrast, *trouble* is more commonly used in fiction and rarely in the academic genre. Additionally, the results from COCA indicated that *problem* typically collocates with verbs of written academic discourse, such as *analyze, eliminate, exacerbate,* etc. Trouble, on the other hand, frequently collocates with lower formality verbs, such as *mean, spell, ask* etc. Jarunwaraphan and Mallikamas (2020) conducted a corpus-based study of near-synonymous nouns: *opportunity* and *chance* by examining their similarities and differences across five genres of COCA. The results indicated that the noun *opportunity* was more frequently used in academic genres but had a lower frequency in fiction genres. In contrast, the noun *chance* was more commonly used in spoken genres and had the least occurrences in academic genres. Furthermore, the top list of their collocates showed that *opportunity* was mostly associated with a formal style, whereas *chance* was associated with an informal style.

Phoocharoensil (2020) investigated the collocational and genre patterns of near-synonymous nouns, *consequence*, *outcome* and *results* in another study. The data were from the COCA. The findings revealed that all three synonyms commonly appear in academic texts, which means they have the highest frequencies in academic genre. In contrast, they have the lowest frequencies in informal genres of fiction, TV and movie subtitles. Kruawong and Phoocharoensil (2022) also examined the genre and collocational patterns of three near-synonymous verbs, *educate, instruct* and *teach* across eight genres. Again, the data came from the COCA. These findings indicated that *teach* was more widely used than *instruct* and *education* in distributional patterns between genres. In other words, *teach* has a higher frequency than the other two verbs among the eight genres. Moreover, the results also showed that all three verbs were more commonly used in formal genres than spoken genres.

Alanazi (2022) investigated near-synonymous verbs, *affect* and *impact* in terms of their colligational profile, semantic preference, and collocational patterns. The data were from the British National Corpus (BNC). The findings revealed their contextual differences, showing that *affect* was more commonly used than *impact* in the written texts. It also showed that *affect* typically has semantic groupings of *intensity/degree/ emphasis, possibility* and *type/specificity*. In contrast, *impact* has two groups of semantic preference of *gradation* and *intensity/emphasis*. Uba and Irudayasamy (2023) examined behavioral profile of English near-synonymous verbs, *increase* and *rise*. The data were from the BNC. The results showed that both verbs typically collocate with subject and object nouns. However, the semantic preference for *increase* was for abstract nouns related to *economy/finance*. In contrast, *rise* had three different kinds of abstract nouns: *human entity, natural environment,* and one more concrete noun related to *natural environment.*

There are several motivations for conducting this study. Firstly, the emergence of the latest version of the COCA, comprising three newly added genres, namely blogs, webpages, and TV and Movie subtitle, and the collocational analysis based on MI scores aims to examine the differences and similarities of these five near-synonyms among the eight genres of the COCA (Davies, 2020). Secondly, Stewart (2010) and Alanazi (2022) claimed that introspection and intuition can provoke a corpus-based study. Being an English as a second language teacher and a non-native speaker of the English language, the selection of this set of near-synonyms triggered my intuition that these English synonymous verbs might be problematic to master. Additionally, as defined above by the three dictionaries, this set of near-synonyms shares the core meaning of 'to get possession or control of something', but there is a lack of comprehensive descriptions of their genre usage, collocational patterns, and semantic preferences in the dictionaries' entries. Furthermore, many studies have proved that a corpus-based study approach is effective in distinguishing near-synonyms' similarities and differences (Alanazi, 2022; Panrat, & Yanasugondha, 2024; Phoocharoensil, 2020; Kruawong & Phoocharoensil, 2022; Putklang et al., 2024; Uba & Irudayasamy, 2023). Hence, this study aims to investigate the genre usage, collocational patterns, and semantic preferences of English synonymous verbs, *grasp, capture, seize, snatch,* and *take*, by using a corpus-based approach. This study seeks to answer the following research questions:

2.7 Research Questions

1. What are the frequency differences among the English near-synonymous verbs **take, capture, seize, snatch,** and **grasp** across the eight genres of the **COCA**?

2. What are the most common noun, and adverb collocates for this set of near-synonyms across the eight genres of the **COCA**?
3. What are the semantic preferences of this set of near-synonyms across the eight genres of the COCA?

3. Methodology

The data for this study were drawn from the Contemporary Corpus of American English (COCA). The corpus is freely accessable online via this link https://www.english-corpora.org/coca/. The corpus is very large and genre-balanced. Phoocharoensil (2020) claimed it is the most widely used English corpus for English language teaching and research. It has more than one billion words. The recently updated version developed by Davies (2020) added three more genres, making it eight genres. The eight genres include academic texts, popular magazines, newspapers, fiction and spoken text, while the newly added genres comprise webpages, blogs, and TV/movie subtitles. The rationale for using the COCA is twofold: firstly, unlike the BNC, which is static, the COCA is annually expanding its size by adding approximately 20 million words (Phoocharoensil, 2020). Secondly, the corpus's data is equally divided among its eight genres, providing more transparent, up-to-date information and making it the largest and most well-developed corpus (Davies, 2020, & Phoocharoensil, 2020).

The present study aimed to answer three research questions. To address question one, we consulted the COCA using its search tool to generate the frequencies and distributional patterns of the five synonymous verbs across the eight genres. We used the 'verb.ALL' facility, enabling the researchers to select only verb word-class, which allowed us to control the lemmatization of verbs. To address research question two, we searched for the most frequent noun and adverb collocates of the near-synonymous English verbs *take, capture, seize, snatch,* and *grasp* across the eight genres of the COCA. Gablasova et al. (2017) and Phoocharoensil (2020) claim that MI determines how words co-occur by having a strong association regarding collocations or by co-occurring by chance. Despite its advantages, MI has some limitations; for example, a collocation with a high MI score might not be representative across many texts (Phoocharoensil, 2020). Some scholars recommend a combination of both a minimum frequency threshold and MI score in determining a strong collocate (Cheng, 2012; Schmitt, 2010; & Kruawong and Phoocharoensil, 2022). Thus, we considered the Mutual Information scores (MI) and frequency in determining the strength of the top fifteen collocates. In response to question three, we categorized each synonymous verb's top fifteen noun and adverb collocates into semantic groups. This indicates how synonymous words share a semantic feature with their collocates (Sinclair, 2004).

4. Results and Discussion

This section will present and discuss the study's results. Table 2 below presents the overall frequencies of grasp, capture, seize, snatch, and take across the eight genres of COCA.

Table 2. Frequency and distribution of synonyms *grasp, capture, seize, snatch* and *take* across genres

	Grasp		*Capture*		*Seize*		*Snatch*		*Take*	
Genre	Freq.	Per mil	Freq.	Per mil	Freq.	Per mil	Freq.	Per mil	Freq.	Per mil
Academic Texts	1964	16.40	4,941	41.25	686	5.73	60	0.50	42,745	356.83
Webpages	2,119	17.05	3,569	28.72	870	7.00	226	1.82	103,120	829.92
Blog	2,192	17.04	3,046	23.68	630	4.90	212	1.65	110, 778	861.33
Magazines	2,077	16.47	4,571	36.25	846	6.71	283	2.24	88,424	701.27
TV/Movies	610	4.76	1,653	12.91	731	5.71	437	3.41	215, 604	1,683.43
Newspaper	1,114	9.15	2,701	22.19	937	7.70	212	1.74	77, 267	634.68
Fiction	2,616	22.11	1,368	11.56	674	5.70	775	6.55	93,791	792.68
Spoken	579	4.59	2,086	16.54	772	6.12	111	0.88	132,267	1,048.61
Total	**13,271**		**23,935**		**6,146**		**2,316**		**863,996**	

Table 2 above shows that, in COCA, *take* is much more commonly used across all eight genres, with a total frequency of 863,996 out of 909,634 total occurrences of all verbs across these genres. The verb *take* has the highest frequency in non-formal spoken and TV/movies genres. Further frequency distribution indicates that the verb *capture* is the second most frequent, with occurrences of 23,935 and its highest frequency in academic texts. The third most frequent verb is *grasp*, which has the highest frequency in the non-formal fiction genre. The fourth most frequent verb, as shown in Table 2 above, is *seize*, with a total frequency of 6,146 across the eight genres, and its highest frequency is in the newspaper genre. Finally, the least frequent verb is *snatch*, with total occurrences of 2,316 and its highest frequency in the fiction genre. This finding aligns with the theory of genre, which suggests that different genres may impose restrictions and expectations that could influence word choice; genre is an interplay between linguistic, professional, and disciplinary practices (Bhatia, 1993; & Swales, 1990). Thus, the results corroborate the assertion of genre theory, which emphasizes the linguistic choices, restrictions, and expectations of different genres. As shown in Table 2 above, the frequencies of the five synonymous verbs are not evenly co-occurring across the eight genres. Additionally, the findings are consistent with previous studies, which found that near-synonyms typically differ in frequencies and distributional patterns across genres (Phoocharoensil, 2020; & Kruawong & Phoocharoensil, 2022).

The following section presents findings on the top fifteen noun and adverb collocates of the synonymous verbs *grasp, capture, seize, snatch,* and *take.* This addresses the second research question. The top fifteen noun collocates of the synonymous verbs, based on MI score and frequency, are in Table 3 below. The table shows that only a few nouns are shared between the synonymous verbs. For example, *opportunity*

is shared only with *grasp* and *seize*; *hand* is commonly found in *grasp* and *snatch*; *arm* is found in *grasp, seize* and *snatch*; and *force* occurs in *capture* and *seize*. Surprisingly, the most commonly used verb *take*, does not share any noun collocates with the other four synonymous verbs. This finding provides insightful descriptions of their noun collocational behaviors. It shows that despite sharing similar meanings, they are not interchangeable in many contexts. This finding is consistent with Sinclair's (1998) claim that their core meaning (paradigmatic level) cannot only define words, but we must also look at the syntagmatic level view of meaning (top-down presentation). It also aligns with the lexical theory of meaning, in which Hoey (2005, p.116) states "words are primed for one or more collocations, semantic associations, and colligations". This finding also corroborates previous studies of near-synonyms in which a set of near-synonyms does not share similar collocates (Alanazi, 2022; Phoocharoensil, 2020, & Uba & Irudayasamy, 2023).

Table 3. Analysis of Noun Collocates

	Grasp			*Capture*			*Seize*			*Snatch*			*Take*		
	Noun Collocate	Freq.	MI Value	Noun Collocate	Freq.	MI Value	Noun Collocate	Freq.	MI Value	Noun Collocate	Freq.	MI Value	Noun Collocate	Freq.	MI Value
1	Hand	934	4.16	Image	1352	4.29	Opportunity	1321	5.84	Hand	236	3.04	Care	67926	4.33
2	Concept	479	5.85	Attention	1254	4.38	Control	796	4.70	Purse	151	7.50	Place	60846	3.07
3	Straw	326	7.98	Camera	1016	4.62	Power	728	3.76	Jaw	114	7.08	Look	36545	4.00
4	Arm	230	3.72	Moment	1000	3.30	Moment	611	4.02	Phone	95	3.03	Step	29402	3.77
5	Meaning	155	4.7	Imagination	903	6.53	Property	407	4.65	Girl	90	2.25	Advantage	22634	4.53
6	Handle	146	6.53	Video	841	3.73	Government	369	2.47	Defeat	83	6.57	Action	21500	2.93
7	Shoulder	126	3.89	Essence	698	6.54	Asset	279	5.53	Bag	75	3.80	Picture	16231	2.78
8	Mind	119	2.31	Data	513	2.41	Police	271	3.02	Arm	66	2.78	Break	13537	3.55
9	Reality	113	3.37	Film	460	2.67	Chance	206	2.96	Baby	62	2.33	Account	12310	3.09
10	Significance	112	5.56	Spirit	427	3.71	Land	194	3.19	Victory	59	4.09	Breath	12040	4.01
11	Finger	99	3.48	Heart	425	2.25	Force	183	2.51	Paper	58	2.56	Risk	10240	2.31
12	Truth	98	2.67	Soldier	410	3.67	Arm	178	2.85	Ball	42	2.67	Responsibility	9966	3.12
13	Ability	82	2.62	Photo	376	2.36	Weapon	173	3.52	Wind	42	3.17	Turn	7661	2.92
14	Opportunity	81	2.30	Force	365	2.07	Authority	163	3.38	Key	42	3.39	Note	7288	2.16
15	Nature	76	2.42	Energy	287	2.08	Evidence	168	2.61	Glass	39	2.79	Shot	6996	2.28

The next step was to analyze the semantic preferences of the five near-synonymous verbs and their noun collocates based on similarities. As defined above, semantic preference is characterized by lexical items commonly used, restricted, and associated with a specific identifiable area or field (Ang et al., 2017; Cheng, 2012; & Kruawong & Phoocharoensil, 2022). In other words, lexical items are typically restricted to certain semantic features due to their semantic relations. The collocational information provides insightful details of the associations of the synonymous verbs and the semantic relationships between the collocates.

Table 4. Semantic preference of noun collocates of *grasp*

	Semantic category	**Example**
1	Body parts	Hand, arm, shoulder, finger
2	Abstract concepts	Concepts, meaning, reality, truth, significance, nature
3	Physical objects	Handle, straw
4	Mental attributes	Mind, ability
5	Opportunities	Opportunity

Table 4 above shows the semantic themes of noun collocates of *grasp*. As seen in the table, five semantic themes emerged. The first theme is BODY PARTS, comprising *hand, arm, shoulder* and *finger.* The second semantic theme is ABSTRACT CONCEPTS, including *concept, meaning, reality, truth, significance,* and *nature*. PHYSICAL OBJECT is the third theme, involving *handle* and *straw*. The fourth theme is MENTAL ATTRIBUTES, including *mind* and *ability*. Finally, the fifth theme is OPPORTUNITY, which has only one collocate, *opportunity.* Despite sharing a noun collocate, arm by *grasp, seize and snatch*, the noun *arm* has different senses, for example, in the below citation from COCA:

Maman grasped her arm and led her to the main salon…

In this co-text, *arm* refers to a part of the body. However, in other co-texts, *arm* refers to a weapon, which we will discuss in the *seize* category. This shows the fluidity of some lexical items.

Table 5. Semantic preference of noun collocates of *capture*

	Semantic category	**Example**
1	Visual representation	Image, camera, photo, video, film
2	Abstract concepts	Imagination, essence, spirit, moment
3	Focus	Attention
4	Data	Data
5	Force	Force, soldier, energy
6	Body part	Heart

Unlike the semantic themes of noun collocates of *grasp*, the semantic preference of noun collocates of *capture* in Table 5 indicates that the semantic themes are categorized into six groups. The first theme is VISUAL REPRESENTATION, comprising *image, camera, video,* and *film*. The second theme is ABSTRACT CONCEPTS, including *imagination, essence, spirit,* and *moment.* FOCUS is the third theme

with only one collocate, *attention*. Similarly, the fourth theme is DATA, which also has only one collocate, *data*. The fifth theme is FORCE, involving *force, energy,* and *soldier*. BODY PART is the sixth theme with only one collocate, *heart*. The comparison of the semantic themes of *grasp* and *capture* noun collocates shows that they only share two semantic themes: ABSTRACT CONCEPTS and BODY PARTS. Below are examples of citations on *video* and *data* from the COCA:

They saw a video of that captured Indian pilot…

So the data may not fully captured public opinion…

The next semantic preference grouping is noun collocates of *seize*.

Table 6. Semantic preference of noun collocates of *seize*

	Semantic category	Example
1	Control and influence	Control, power, authority
2	Government and law enforcement	Government, police, force
3	Asset	Property, asset, land
4	Abstract concept	Moment,
5	Opportunity	Opportunity, chance
6	Weapon	Weapon, arm
7.	Proof	Evidence

As seen in Table 6, the semantic preferences of noun collocates of *seize* were classified into seven semantic themes. The first theme is CONTROL/INFLUENCE, comprising *control, power,* and *authority*. The second theme is GOVERNMENT/LAW ENFORCEMENT, including *government, police,* and *force*. ASSET is the third theme, involving *property, asset,* and *land*. The fourth theme is ABSTRACT CONCEPT, with only one collocate, moment. The fifth theme is OPPORTUNITIES, comprising *opportunity* and *chance*. The sixth theme is WEAPON, including *weapon and arm*. The last theme PROOF has only one collocate, *evidence.* As noted above, some lexical items have multiple senses; here, *arm* has a different meaning from the previous one discussed in Table 4 above. Here the meaning refers to the ammunition, whereas in Table 4 it refers to a part of the body. An example from the COCA is:

We must arm the insurgents. Arm them and train them…

This aligns with the theory of meaning, which emphasizes that the meaning of a word can only be understood or differentiated by not only looking at its core meaning but also the context/co-text of words which usually occur (Firth, 1957; Sinclair, 1998; & Hoey, 2005).

Table 6. Semantic preference of noun collocates of *snatch*

	Semantic category	Example
1	Body parts	Hand, jaw, arm
2	Physical object	Purse, phone, bag, paper, key, glass, ball
3	People	Girl, baby
4	Abstract concepts	Defeat, victory
5	Nature	Wind

Table 6 shows the semantic preference grouping of noun collocates of *snatch*. The table indicates that BODY PART is the first semantic theme, including the noun collocates *hand, jaw,* and *arm.* The second semantic theme is PHYSICAL OBJECT, comprising the noun collocates *purse, phone, bag, paper, key, glass,* and *ball*. The third theme is PEOPLE, including the noun collocates *girl* and *baby*. The fourth theme is ABSTRACT CONCEPTS, involving the noun collocates *defeat* and *victory*. The final theme is NATURE, with only one noun collocate, *wind*. Below are two examples of *wind* and *bag* provided from the COCA:

The north wind snatches the handle away from her…

She snatched her bag and marched to the door...

The following table is a semantic preference groupings of noun collocates of *take.*

Table 7. Semantic preference of noun collocates of *take*

	Semantic category	Example
1	Responsibility	Care, responsibility, risk
2	Actions	Action, step, turn, break
3	Observation and perception	Look, note, picture, shot
4	Benefit	Advantage
5	Life	Breath
6	Location	Place
7	Evaluation	Account

Table 7 indicates the semantic preference grouping of nouns collocating with *take.* As seen in the table, there are seven semantic groupings of noun collocates of *take*. The first theme is RESPONSIBILITY, which includes nouns that collocate *care, responsibility,* and *risk.* ACTION is the second semantic theme, involving noun collocates *action, step, turn,* and *break*. The third semantic theme is OBSERVATION, comprising noun collocates *look, note, picture,* and *shot*. The fourth semantic theme is BENEFIT, with only one collocate, *advantage*. LIFE is the fifth theme and has only one noun collocate: *breath*. LOCATION is the sixth theme, with one noun collocate, *place*. Finally, EVALUATION has one noun collocate: *account*. Some examples from the COCA on *care* and *place* noun

collocates are provided below:

But great care needs to be take…

There are so many places I can't take my dog on…

Analyzing the semantic preferences of noun collocates of the five near-synonymous verbs indicates that some verbs share four different semantic themes. For example, the semantic theme BODY PART is shared by *grasp, capture and snatch*. The second theme shared by *grasp*, *capture*, *seize*, and *snatch* is ABSTRACT CONCEPT. OPPORTUNITY is the third theme shared by *seize* and *grasp* only. Finally, PHYSICAL OBJECT is also shared by only *snatch and grasp*. Surprisingly, the highest frequency verb, *take* does not share any of its semantic themes with the other four verbs, nor do the four verbs share their semantic themes with the verb *take*. This finding is consistent with genre theory, as a corpus-based study could reveal subtle differences among near-synonyms in terms of collocation, formality, semantic preferences, and connotation by showing their appropriateness in certain genres (Bhatia, 1992; & Swales, 1974). Additionally, the results corroborate previous studies, such as Alanazi (2022), which found that *affect* typically has semantic groupings of *intensity/degree/ emphasis, possibility* and *type/specificity*. In contrast, *impact* has two groups of semantic preference of *gradation* and *intensity/emphasis*. Similarly, Uba and Irudayasamy (2023) revealed that the semantic preference for *increase* was for abstract nouns related to *economy and finance*. In contrast, *rise* had three different kinds of abstract nouns, related to *human entity, natural environment*, and one more concrete noun related to *natural environment*. This clearly shows that despite the target near-synonyms sharing some core meanings, they are not interchangeable in many contexts.

The top fifteen adverb collocates of the synonymous verbs based on MI score and frequency are in Table 8 below. The table shows that only nine adverbs are shared between the synonymous verbs. For example, *fully* is shared by *grasp* and *capture*; *quickly* is found in *grasp*, *seize*, and *snatch; immediately* is found in *grasp* and *seize*; *allegedly* is commonly shared with *seize and snatch*; *firmly* is found in *grasp* and *seize*; and *eagerly* occurs in *seize* and *snatch*. Moreover, *instantly* is shared by *grasp* and *seize*; *suddenly* is found in *seize* and *snatch*; and finally, *away* is found in *snatch* and *take*. This shows that the highest frequency verb, *take*, shared only one collocate with *snatch*.

Table 8. Analysis of Adverb Collocates

	Grasp			*Capture*			*Seize*			*Snatch*			*Take*		
	Adverb Collocate	Freq	MI Value	Adverb Collocate	Freq.	MI Value	Adverb Collocate	Freq.	MI Value	Adverb Collocate	Freq.	MI Value	Adverb Collocate	Freq.	MI Value
1	Fully	363	5.72	Perfectly	260	4.05	Immediately	84	2.78	Up	1395	3.48	On	47235	2.02
2	Quite	152	2.95	Best	199	2.29	Suddenly	84	2.82	Away	526	4.64	Off	43963	2.69
3	Quickly	121	3.21	Fully	162	2.63	Quickly	77	2.07	Quickly	47	2.72	Away	26898	2.41
4	Firmly	99	6.04	Accurately	123	4.63	Illegally	51	5.36	Suddenly	43	3.22	Long	20737	2.07
5	Easily	70	3.08	Adequately	74	4.66	Eagerly	25	5.16	Forward	37	2.36	Seriously	18518	4.33
6	Hard	62	2.14	Beautifully	71	4.36	Violently	17	4.59	Behind	11	2.21	Aback	1805	6.11
7	Tightly	60	5.57	Nicely	57	3.81	Temporarily	13	3.12	Abruptly	10	4.41	Lightly	1236	2.49
8	Immediately	58	2.74	Vividly	41	4.99	Briefly	12	2.02	Allegedly	9	3.41	Awhile	825	2.91
9	Intuitively	57	8.37	Brilliantly	40	4.96	Instantly	11	2.31	Eagerly	8	4.88	Kindly	527	3.19
10	Truly	53	2.85	Successfully	38	2.06	Allegedly	11	2.34	Hastily	7	5.25	Gladly	244	2.14
11	Gently	33	3.86	Realistically	32	4.62	Forcibly	11	5.09	Secretly	5	3.34	Offline	133	2.54
12	Readily	33	4.50	Succinctly	29	5.74	Legally	10	2.10	Quick	5	3.70	Forcibly	130	2.11
13	Desperately	33	4.60	Neatly	23	3.07	Thereby	10	2.10	Angrily	5	4.58	Orally	125	2.77
14	Barely	25	2.45	Sufficiently	22	2.65	Firmly	10	2.24	Round	4	2.61	Guard	109	2.21
15	Instantly	16	3.34	Wonderfully	19	3.46	Reportedly	9	2.03	Promptly	4	3.44			

The next step was to analyze the semantic preferences of adverb collocates for the five near-synonymous verbs.

Table 8. Semantic preference of adverb collocates of *grasp*

	Semantic category	**Example**
1	Degree or Extent (how much)	Quite, fully barely, truly,
2	Manner of action	Quickly, firmly, tightly, gently, easily, readily, hard
3	Time and speed	Immediately, Instantly, desperately
4	Intuition	Intuitively

Table 8 shows the semantic themes of adverb collocates of *grasp*. As seen in the table, four semantic themes emerged. The first theme is DEGREE/EXTENT, which includes the adverb collocates *fully, quite, barely, and truly*. The second semantic theme is MANNER OF ACTION, comprising the adverb collocates *quickly, firmly, tightly, gently, easily, readily,* and *hard*. The third semantic theme is TIME/SPEED, involving the adverb collocates *immediately, instantly,* and *desperately*. INTUITION is the last semantic theme, with only one adverb collocate, *intuitively*.

Table 9. Semantic preference of adverb collocates of *capture*

	Semantic category	Example
1	Degree of completion	Fully, adequately, sufficiently
2	Precision	Perfectly, accurately, realistically, succinctly
3	Achievement	Best, successfully
4	Aesthetic quality	Wonderfully, brilliantly, vividly, nicely, beautifully
5	Neatness	Neatly

The semantic preference in Table 9 indicates five semantic themes of adverb collocates for *capture*. The first theme is DEGREE OF COMPLETION, comprising the adverb collocates of *fully, adequately,* and *sufficiently*. The second theme is PRECISION, including the adverb collocates of *perfectly, accurately, realistically, and succinctly*. ACHIEVEMNET is the third semantic theme, involving the adverb collocates of *best and successfully*. The fourth semantic theme is AESTHETIC QUALITY, comprising adverb collocates of *wonderfully, brilliantly, vividly, nicely,* and *beautifully*. The last theme is NEATNESS, which has only one adverb collocate, *neatly*. As stated earlier, some lexical items may appear in multiple semantic groups because of the co-texts and contexts. Here, the adverb *fully* is under *degree of completion* and in the *grasp* category, it falls under *degree or extent*, that is, how much.

Table 10. Semantic preference of adverb collocates of *Seize*

	Semantic category	Example
1	Time and speed	Immediately, suddenly, quickly, instantly, briefly, temporarily
2	Legitimacy	Illegally, legally, reportedly, allegedly
3	Force	Violently, forcibly
4	Manner of action	Firmly
5	Motivational state	Eagerly
6	Causality	Thereby

The semantic preference in Table 10 shows six semantic themes of adverb collocates for *seize*. The first theme is TIME/SPEED, including the adverb collocates of *immediately, suddenly, quickly, instantly, briefly,* and *temporarily*. The second semantic theme is LEGITIMACY, involving the adverb collocates of *illegally, legally, reportedly,* and *allegedly*. The third theme is FORCE, comprising the adverb collocates of *violently* and *forcibly*. MANNER OF ACTION is the fourth theme, involving the adverb collocate of *firmly.* The fifth theme is MOTIVATIONAL STATE, which has only one adverb collocate, *eagerly*. CAUSALITY is the last theme, which also has only one adverb collocate, *thereby.*

Table 11. Semantic preference of adverb collocates of *snatch*

	Semantic category	Example
1	Time and speed	Suddenly, quickly, abruptly, hastily, promptly, quick
2	Direction or position	Up, away, forward, behind, round
3	Motivational state	Eagerly
4	Emotional state	Angrily
5	Secrecy	Secretly
6	Legitimacy	Allegedly

Table 11 above shows the semantic themes of adverb collocates of *snatch*. As seen in the table, six semantic themes emerged. The first theme is TIME/SPEED, including the adverb collocates *suddenly, quickly, abruptly, hastily, promptly,* and *quick*. The second theme is DIRECTION/POSITION, involving the adverb collocates of *up, away, forward, behind* and *around*. The third theme is MOTIVATIONAL STATE, which has only one collocate, *eagerly*. The fourth theme is EMOTIONAL STATE, which has only one collocate, *angrily.* Furthermore, the fifth theme is SECRECY, having only one collocate, *secretly*. Finally, LEGITIMACY has only one adverb collocate, *allegedly*.

Table 12. Semantic preference of adverb collocates of *take*

	Semantic category	Example
1	State or condition	On, off, away, offline
2	Duration	Long, awhile
3	Manner or intensity	Seriously, lightly, forcibly, orally, kindly, gladly
4	Emotional state	Angrily
5	Secrecy	Secretly
6	Legitimacy	Allegedly
7	Reactions or surprises	Aback
8	Defense	guard

The semantic preferences of the most frequent verb *take* for adverb collocates are shown in Table 12. As seen in the table, eight semantic themes emerged. The first theme is STATE/CONDITION, including adverb collocates *on, off, away,* and *offline*. The second theme is DURATION, including adverb collocates *long* and *awhile*. The third theme is MANNER/INTENSITY, comprising *seriously, lightly, forcibly, orally, kindly,* and *gladly*. The fourth theme is EMOTIONAL STATE, which has only one collocate, *angrily*. The fifth theme is SECRECY, which also has one collocate, *secretly*. The sixth theme is LEGITIMACY, which also involves one collocate, *allegedly*. REACTION/SURPRISE is the seventh theme, which also has one collocate, *aback*. Finally, DEFENSE has one collocate, *guard.*

Unlike the analyzis of the semantic preferences of noun collocates for the five near-synonymous verbs, this analysis of the semantic preferences of the adverb collocates shows that the four synonyms share six semantic themes. For example, SPEED/TIME is shared by *grasp, seize,* and *snatch*. MANNER OF ACTION is shared only by *grasp* and *seize*. LEGITIMACY is shared by *seize, snatch,* and *take*. MOTIVATIONAL STATE has appeared in both *seize* and *snatch.* EMOTIONAL STATE has occurred in *take, snatch,* and *seize*. Finally, SECRECY is shared by *snatch* and *take*. This clearly indicates that the verb *capture* does not share any of its semantic themes with the other four verbs. Again, this analysis supports the claim that 'natural languages abhor absolute synonyms just as nature abhors a vacuum' (Cruse, 1986, p. 270). Here, despite sharing the core meaning, the target synonym shows that they are not absolute synonyms. Additionally, Flowerdew (2012) claims that the meaning of words does not solely depend on their inherent qualities but also includes how they collocate with other words. In other words, synonyms can best be examined by looking at their collocates to determine similarities and differences. The previous studies also align with this finding. For example, Gu (2017) found that *gain* is associated with abstract nouns and *obtain* commonly occurs with a noun in the passive voice. Phoocharoensil (2020) also revealed that *consequence* is associated with negative senses, *outcome* has the broadest senses, and *result* is associated with academic contexts.

This study makes five key contributions to the existing literature. Firstly, it aligns with the claim that traditional reference materials, such as dictionaries, do not provide a comprehensive description of near-synonyms, including genre usage, collocational behavioral profiles, and semantic preferences (Alanazi, 2022; Panrat & Yanasugondha, 2024; Phoocharoensil, 2020; Kruawong & Phoocharoensil, 2022; Putklang, et al., 2024; Uba & Irudayasamy, 2023). However, this study, and indeed previous studies, prove that a corpus-based study of near-synonyms provides comprehensive descriptions of genre usage, collocational behavioral profiles, and semantic preferences (Hunston, 2002; Liu, 2010; Jirananthiporn, 2018; & Kruawong & Phoocharoensil, 2022). This study shows that the five synonymous verbs collocate with different nouns and adverbs. They typically have different frequencies across the eight genres. Additionally, they do not share all the semantic preferences, implying that they cannot be substituted in many contexts.

Secondly, the findings of this study are consistent with a key principle of lexical theory in that the meaning of a word does not solely depend on what Sinclair (1998) calls the 'paradigmatic level' (core meaning), but also involves the 'syntagmatic level' (top-down level). Additionally, Hoey (2005, p. 116) states that 'words are primed for one or more collocations, semantic associations …'. This study illustrates how the analyzes of genre usage, collocational and semantic preferences of the target near-synonyms provide insightful descriptions and knowledge of their behavioral profiles. The analyzes shows that each verb typically co-occurs with certain nouns and adverbs. Additionally, each verb is typically associated with or has semantic relationships with certain lexical items. This finding corroborates previous studies of corpus-based analysis on near-synonyms (Phoocharoensil, 2020; Kruawong & Phoocharoensil, 2022; Putklang, et al., 2024; Uba & Irudayasamy, 2023).

Thirdly, the finding aligns with key principles of genre theory, which emphasizes that different genres may impose restrictions and expectations that influence word choice. Moreover, a corpus-based genre and collocational near-synonyms study could reveal subtle differences in formality, connotation, semantic preference, and collocation, by demonstrating their suitability for certain genres (Swales, 1990 & Bhatia, 2004). As shown in Table 2, there are different overall frequencies of the target synonymous verbs across eight genres in COCA. This clearly demonstrates how genres impose restrictions and expectations that could influence word choice. For example, the verb *take* has a low frequency in the academic genre but has the highest frequency in the non-formal genre of TV/Movies. This finding also aligns with previous studies on genre and collocational analysis of near-synonyms (Phoocharoensil, 2020 & Kruawong & Phoocharoensil, 2022).

Fourthly, this finding could assist teachers in raising the students' awareness of collocational patterns, genre usage, and semantic preferences of this set of near-synonyms. It indicates that this set of near-synonyms is commonly associated with particular words, allowing language learners to acquire and learn different contexts of use for such synonyms. For example, teachers could ask students to generate keywords in the context of 20 citations (concordance) and identify noun collocates of the target synonyms. Additionally, the COCA is freely accessible online. Teachers could engage students in task-based learning. For example, they can ask students to generate a collocational distribution list of a set of near-synonyms and then group the semantic preferences of the target synonyms. The students could find subtle differences between the synonyms.

Finally, this study addresses significant gaps in previous research by offering context-sensitive, genre-sensitive, and empirical insights beyond simplistic definitions and native speakers' intuitions. While earlier studies often overlooked genre variations and relied on introspection, this approach utilizes a large English language corpus to demonstrate how this set of near-synonyms varies regarding grammatical patterns, frequency, collocation, and usage across different genres. It provides practical insights for language teaching and lexicography by promoting a more comprehensive understanding of synonym behaviour.

5. Conclusion

This corpus-based genre and collocational study of the near-synonym verbs, *grasp, capture, seize, snatch* and *take* provides comprehensive descriptions of their overall frequencies, collocational behavioural profile and semantic preferences across eight genres of COCA. The results provide insightful information associated with their noun and adverb collocates in authentic texts. This could benefit learners more than relying on traditional non-corpus-based reference materials such as dictionaries or native speakers' intuition. It also yielded information on their semantic preferences, where we found that the target synonyms have specific semantic relationships with certain noun or adverb collocates, which other synonyms do not possess. For example, the verb *capture* does not share any semantic

theme with its adverb collocates with the other four verbs. Similarly, the highest frequency verb, *take*, does not share any of its semantic themes with the other four verbs, nor do the four verbs share their semantic themes with the verb *take*.

One limitation of this study was the involvement of only one corpus, COCA. Future studies could include more corpora for triangulation of the results. Moreover, this study could not include adjective collocates of the target synonyms due to word limit constraints. We recommend that future studies incorporate adjective collocates. As noted by Schmitt (2010), collocations generated based on a particular type of statistical measure might differ. Future studies could use different statistical measures, such as log-likelihood, t-score, and z-score. Additionally, we restricted the number of collocates to 15; future studies could expand the range to more than 15.

Acknowledgments

Not applicable.

Authors' contributions

Dr. SYU was responsible for the study design and literature review. Dr. SYU and Dr. FA were responsible for the research design, data analysis, and presentation. Dr. SYU prepared the discussion and conclusions, while Dr. FA was responsible for critically reviewing, proofreading, and editing the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Not applicable.

Ethics approval

The Publication Ethics Committee of the Sciedu Press.

The journal's policies adhere to the Core Practices the Committee on Publication Ethics (COPE) established.

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data supporting this study's findings are available on request from the corresponding author. However, due to privacy or ethical restrictions, the data are not publicly available.

Data sharing statement

No additional data are available.

References

Alanazi, Z. (2022). Corpus-based analysis of near-synonymous verbs. *Asian-Pacific Journal of Second and Foreign Language Education*, *7*(1), 15. https://doi.org/10.1186/s40862-022-00138-5

Ang, L. H., Tan, K. H., & He, M. (2017). A corpus-based collocational analysis of noun premodification types in academic writing. *3L, Language, Linguistics, Literature*, *23*(1). http://doi.org/10.17576/3L-2017-2301-09

Bhatia, V. K. (1992). Pragmatics of the Use of Nominals in Academic and Professional Genres. *Pragmatics and language learning*, *3*, 217-230.

Bhatia, V. K. (1993). *Analyzing genre: Language use in professional settings.* Longman.

Bhatia, V. K. (2004). *Integrating Content and Language: Meeting the Challenge of a Multilingual Higher Education.* Maastricht: Universitaire Pers Maastricht.

Chan, A. Y. (2010). Toward a taxonomy of written errors: Investigation into the written errors of Hong Kong Cantonese ESL learners. *Tesol Quarterly*, *44*(2), 295-319. https://doi.org/10.5054/tq.2010.219941

Cheng, W. (2012). *Exploring corpus linguistics: Language in action*. Routledge. https://doi.org/10.4324/9780203802632

Church, K., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational linguistics*, *16*(1), 22-29.

Cruse, D. A. (1986). *Lexical semantics*. Cambridge University Press.

Davies, M. (2020). The COCA corpus (new version released March 2020). *English Corpora*. Retrieved from https://www.english-corpora.org/coca/

Divjak, D. (2006). Ways of intending: Delineating and structuring near-synonyms. *Trends in Linguistics Studies and Monographs*, *172*, 19. https://doi.org/10.1515/9783110197709.19

Firth, J. (1957). *Papers in Linguistics.* Oxford University Press

Flowerdew, L. (2012). *Corpora and language education*. Palgrave Macmillan. https://doi.org/10.1057/9780230355569

Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language learning*, *67*(S1), 155-179. https://doi.org/10.1111/lang.12225

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied linguistics*, *35*(3), 305-327. https://doi.org/10.1093/applin/amt015

Gu, B. J. (2017). Corpus-based study of two synonyms: obtain and gain. *Sino-US English Teaching*, *14*(8), 511-522. https://doi.org/10.17265/1539-8072/2017.08.006

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Routledge.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524773

Jackson, H., & Amvela, E. Z. (2007). *Words, meaning and vocabulary: An introduction to modern English lexicology*. Bloomsbury Publishing. https://doi.org/10.5040/9781350934047

Jarunwaraphan, B., & Mallikamas, P. (2020). A corpus-based study of English synonyms: Chance and opportunity. *rEFLections*, *27*(2), 218-245. https://doi.org/10.61508/refl.v27i2.248710

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.

Kruawong, T., & Phoocharoensil, S. (2022). A genre and collocational analysis of the near-synonyms teach, educate and instruct: A corpus-based approach. *TEFLIN*, *33*(1). https://doi.org/10.15639/teflinjournal.v33i1/75-97

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*(1), 56-87. https://doi.org/10.1075/ijcl.15.1.03liu

Longman dictionary of contemporary English online. Retrieved from https://www.ldoceonline.com/dictionary/take

Meriam-Webster dictionary. Retrieved from https://www.merriam-webster.com/

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139858656

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, *24*(2), 223-242. https://doi.org/10.1093/applin/24.2.223

Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research*, *21*(3), 298-320. https://doi.org/10.1177/1362168816639619

Oxford Advanced American Dictionary. https://www.oxfordlearnersdictionaries.com/definition/american_english

Panrat, T., & Yanasugondha, V. (2024). A Corpus-Based Study of English Synonyms" Clear,"" Obvious,"" Apparent," and" Evident": Implications for ELT. *LEARN Journal: Language Education and Acquisition Research Network*, *17*(1), 162-187. Retrieved from https://so04.tci-thaijo.org/index.php/LEARN/index

Partington, A. (1998). *Patterns and meanings: Using corpora for English language research and teaching*. John Benjamins. https://doi.org/10.1075/scl.2

Phoocharoensil, S. (2020). A Genre and Collocational Analysis of Consequence, Result, and Outcome. *3L: Southeast Asian Journal of English Language Studies*, *26*(3). http://doi.org/10.17576/3L-2020-2603-01

Phoocharoensil, S. (2021). Semantic prosody and collocation: A corpus study of the near-synonyms persist and persevere. *Eurasian journal of applied linguistics*, *7*(1), 240-258. https://doi.org/10.32601/ejal.911269

Putklang, T., Panthong, P., & Chutopama, N. (2024). A Corpus-based Analysis of Academic Near Synonyms: Consist, Comprise and Compose. *The Golden Teak: Humanity and Social Science Journal*, *30*(4), 129-148.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Springer. https://doi.org/10.1057/9780230293977

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.

Sinclair, J. (1998). *The lexical item*. John Binjamins B.V.. https://doi.org/10.1075/cilt.171.02sin

Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. Routledge. https://doi.org/10.4324/9780203594070

Stewart, D. (2010). *Semantic prosody: A critical evaluation*. Routledge. https://doi.org/10.4324/9780203870075

Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of language*, *2*(1), 23-55. https://doi.org/10.1075/fol.2.1.03stu

Swales, J. M. (1974). Notes on the function of attributive en-participles in scientific discourse. Papers for special university purpose No. 1 ELSU, University of Khartoum

Swales, J. M. (1990). *Genre analysis*. Cambridge university press.

Szudarski, P. (2018). *Corpus linguistics for vocabulary: A guide for research: Routledge.* https://doi.org/10.4324/9781315107769

Thornbury, S. (2002). *How to teach vocabulary*. Harlow.

Uba, S. Y., & Irudayasamy, J. (2023). Is It'Increase'or'Rise?'A Corpus-Based Behavioural Profile Study of English Near-Synonym Verbs. *Mextesol Journal*, *47*(1), n1. https://doi.org/10.61871/mj.v47n1-14

Widdowson, H. (2007). Jr firth, 1957, papers in linguistics 1934–51. *International Journal of Applied Linguistics*, *17*(3), 402-413. https://doi.org/10.1111/j.1473-4192.2007.00164